现代数学基础

24

变分学讲义

BIANFENXUE JIANGYI

■ 张恭庆 编著


高等教育出版社·北京
HIGHER EDUCATION PRESS BEIJING

内容简介

变分学是数学分析的一个重要组成部分，是一门与其他数学分支密切联系、并有广泛应用的数学学科。近几十年来，变分学不论是在理论上还是在应用中都有了很大发展，与数学其他分支的联系也更加紧密，已经成为大学数学教育不可缺少的部分。

本书是作者在北京大学为高年级本科生和低年级研究生开设“变分学”课程所用的讲义。全书共二十讲，分为三大部分：第一部分（一到八讲）是经典变分学的基本内容，第二部分（九到十四讲）重点介绍直接方法及其理论基础，第三部分（十五到二十讲）是专题选讲。其材料的选取，内容的编排，问题与概念的表述，以及证明的分析与讲解均极具特色。

本书适用于数学及相关专业的本科生、研究生、教师以及研究人员，也可供工科、经济学、管理学等专业的教师和学生使用参考。

图书在版编目（CIP）数据

变分学讲义 / 张恭庆编著. —北京：高等教育出版社，2011.6

ISBN 978-7-04-031958-3

Ⅰ. ①变… Ⅱ. ①张… Ⅲ. ①变分学—高等学校—教材 Ⅳ. ① O176

中国版本图书馆 CIP 数据核字（2011）第 094078 号

策划编辑 赵天夫　　责任编辑 赵天夫　　封面设计 赵 阳　　责任印制 张泽业

出版发行 高等教育出版社
社　　址 北京市西城区德外大街 4 号
邮政编码 100120
印　　刷 中国农业出版社印刷厂
开　　本 787×1092 1/16
印　　张 20.75
字　　数 350 000
购书热线 010-58581118
咨询电话 400-810-0598
网　　址 http://www.hep.edu.cn
　　　　 http://www.hep.com.cn
网上订购 http://www.landraco.com
　　　　 http://www.landraco.com.cn
版　　次 2011 年 6 月第 1 版
印　　次 2011 年 6 月第 1 次印刷
定　　价 49.00 元

物 料 号 31958-00

前言

变分学 (calculus of variations) 差不多和微积分同时诞生, 至今已有三百多年的历史. 它曾是大学数学系本科的必修课程, 安排在学完微积分和常微分方程之后讲授, 主要内容是把变分问题的求解化归为微分方程. 然而只有很少几类常微分方程能够把解显式地写出来, 因此能够深入研究的变分问题是很有限的. 过去半个多世纪, 这门课在许多学校逐渐地被削减、合并, 直至分散到其他课程中去了.

然而变分学与许多数学分支有紧密的联系, 在 Hilbert 的 23 个问题中就有 3 个问题属于变分问题, 足以表明其重要性. 变分问题又直接来源于力学、物理、经济、控制、工程等学科. 特别是自 20 世纪 70 年代以来, 有限元方法和最优化技术使得变分极值问题得以数值求解, 明显地提高了变分学在应用数学中的地位.

近几十年来, 变分学不论在理论上还是在应用中都有了重大的发展. 国内外数学界都已经注意到在数学系本科教学中没有变分学不能适应学科发展的要求. 至于怎样弥补不足, 则尚在探索之中. 本书就是在这种背景下的一种尝试.

作者分别于 2006 年与 2010 年在北京大学数学科学学院为高年级本科生和低年级研究生讲授了一门 "变分学" 课程, 并按照下列三条原则来组织教学内容:

1. 既要讲经典理论, 也要讲近代变分理论和它的现代发展, 根据研究问题的不同层次逐步深入.
2. 最经常用到的理论和方法作为重点.
3. 面向全院各专业 (包括基础数学、计算数学、统计数学、信息数学与金融数学) 的学生.

我们要求听课的学生除了学过 "三高" (数学分析、高等代数、解析几何) 外, 还需

具备常微分方程、实变函数、泛函分析、微分几何、数学物理方程等方面的预备知识.

整个课程分三个阶段: 经典变分学、解的存在性与正则性、专题选讲, 后者涉及变分学的现代发展. 本书不但介绍变分学的基本概念、基本理论和基本方法, 而且还通过大量列举正例和反例来阐述这些内容: 界定新概念的内涵与所建立理论和方法的适用范围. 在前两个阶段的每讲之后附有习题, 帮助巩固学过的内容.

本书是在授课讲义的基础上整理编写而成的.

由于这门课程的开设尚在探索之中, 其内容之选取、深浅之掌握都不成熟, 加之作者学识所限, 疏漏、不足之处在所难免, 热忱欢迎读者批评指正.

作者感谢高等教育出版社的赵天夫编辑和陆珊年编辑. 他们细心的阅读和校对以及提出的宝贵意见为本书增色不少.

张恭庆　谨识

2010 年 12 月于北京大学

目录

第一讲　变分学与变分问题

§1.1　前言

变分学是数学分析的一个重要组成部分, 是一门与其他数学分支密切联系、并有广泛应用的数学学科. 例如:

◇ 大量重要的数学物理方程, 弹性、塑性力学中的微分方程, 生物薄膜方程, 几何中的微分方程等都是泛函的 Euler 方程.

◇ 最优控制理论问题是一类与传统不同的带有约束的变分问题, 在工程控制和经济学理论中经常出现.

此外, 智能材料、图像处理、最优工程设计中也出现大量新型的变分问题.

◇ 变分方法是证明椭圆型偏微分方程解的存在性的主要方法, 它已成为偏微分方程理论的重要组成部分. 变分学与偏微分方程关系之紧密可以从 Hilbert 第 19 问题与第 20 问题中看出.

◇ 偏微分方程的数值方法, 特别是有限元方法直接来自变分问题. 最优化技术的发展使得变分极值问题可以直接数值求解.

◇ 拓扑学与变分学的结合产生了一个新的分支 —— 大范围变分学, 也带来了临界点理论的大发展. 特别是 Morse 理论建立了分析与拓扑学的联系, 成为微分拓扑学的一个重要部分. Floer 同调也是这二者结合的产物.

◇ 变分理论深入到 Riemann 几何、Finsler 几何、辛几何、保形几何等学科之中. 一些几何变分问题如测地线、极小曲面、调和映射等的研究更激发了许多

对新理论 (如几何测度论)、新方法和新技巧的研究.

◇ 变分方法在 Hamilton 动力系统的周期轨道、Mather 集、混沌等的研究中起重要作用.

◇ 作为微分学与概率论合并之后产生的随机变分学 (Malliavin Calculus) 成为金融数学的重要部分.

由此可见, 变分学的问题、理论、方法已深入到许多近代数学包括基础数学、应用数学、计算数学、信息数学、经济数学等领域, 它在近代数学中已经占据了重要的地位.

本书与传统的教科书 (例如 [LL], [El], [Ka], [GF]) 相比, 有以下特点:

◇ 在经典变分学部分重点突出了极小点的一阶条件和二阶条件.

因为偏微分方程、微分几何、数学物理中大量有用的例子都是多个变量的, 所以对多个变量的情形花费了一定的篇幅展开讨论.

◇ 在经典理论部分增强了 Hamilton-Jacobi 理论与守恒律两节的内容, 因为它们在物理学与几何学中都是极为重要的.

◇ 除经典的变分理论外, 我们着重介绍了直接方法与应用. 直接方法是近代变分理论的重要组成部分, 也是微分方程解的存在性证明以及数值计算的基础. 这部分内容对学过泛函分析初步的学生是能够接受的. 它占据整个课程将近一半的篇幅.

◇ 特征值问题是分析数学的核心内容之一, 我们把它作为约束极值问题存在性理论的应用来介绍, 正好与泛函分析相应内容相呼应.

此外, 我们还对几个理论与应用问题做了专题介绍, 它们可以作为选讲的内容.

◇ 临界点理论是近几十年来变分学发展最快的一个分支, 也有许多应用. 特别是在微分方程解的存在性的证明中它是直接方法的一个重要补充. 这个理论的内容非常丰富, 我们只可能介绍其中最简单的一个定理 —— 山路定理, 作为临界点理论的入门.

◇ Hamilton 系统的周期解、同宿轨、异宿轨问题是动力系统与辛几何的热点问题, 在一定条件下可以用变分方法证明它们的解的存在性.

◇ 测地线与极小曲面是几何变分问题的最简单的例子. 这一讲也可以看成是几何分析的一个导引.

◇ 变分问题的数值求解主要用有限元与数值优化技术, 而有限元方法的理论

又是建立在变分理论基础之上的.

◇ 可以根据情况, 再选择一个有实际背景的问题, 如最优控制问题、图像处理问题, 以了解变分问题新的理论与应用. 这几讲的内容是相互独立的.

全书共分二十讲. 前八讲是经典的变分学, 第九讲到第十四讲介绍直接方法, 这两部分是本书基本的内容; 从第十五讲到第二十讲是一些专题介绍, 可以作为选讲的材料. 书中带星号 * 的章节, 初读时可以略去.

§1.2 泛函

变分学是研究泛函极值 (以及更一般的临界值) 的一个数学分支.

一般地, 人们把从任意集合 M 到实数域 $\mathbb{R}$ 或复数域 $\mathbb{C}$ 的映射通称为泛函. 但在变分学中, 泛函只取实值, 定义域 M 是一个函数集合, 即 $I: M \to \mathbb{R}$.

例如, 设 $\Omega \subset \mathbb{R}^n$ 是一个有界开集, $x_0 \in \Omega$ 是一个固定点, $F \in C(\bar{\Omega})$, $M = C^1(\bar{\Omega})$.

$$I_1(u) = \max_{x\in\bar{\Omega}} |u(x)|, \quad I_2(u) = u(x_0), \quad I_3(u) = \int_\Omega [|\nabla u(x)|^2 - F(u(x))]dx$$

都是泛函. 但不论取什么样的 M 和一元函数 f, 复合函数

$$I_4(u) = f(u(x))$$

都不是泛函!

给定一个函数 $L \in C^1(\bar{\Omega} \times \mathbb{R}^N \times \mathbb{R}^{nN})$, 变分学主要研究如下形式的泛函:

$$I(u) = \int_\Omega L(x, u(x), \nabla u(x))dx,$$

其中 M 是连续可微函数类 $C^1(\bar{\Omega})$ 的子集合, 或者是某种广义可微函数类的子集合, 它由一些限制条件如 (积分形式的、逐点的、带微分的、不带微分的) 边值条件和约束条件等所规定.

有时, I 的积分中还可以含有高阶导数项, M 也应做相应的改变.

§1.3 典型例子

例 1.1 (最速下降线) 在垂直平面上给定两点 $A = (x_1, y_1)$ 和 $B = (x_2, y_2)$, 其中 $x_1 < x_2$, $y_1 > y_2$. 一个质点沿着一条连接这两点的光滑曲线仅借重力下滑. 设初速度为零, 问沿怎样的一条曲线滑行时间最短 (见图 1.1)?

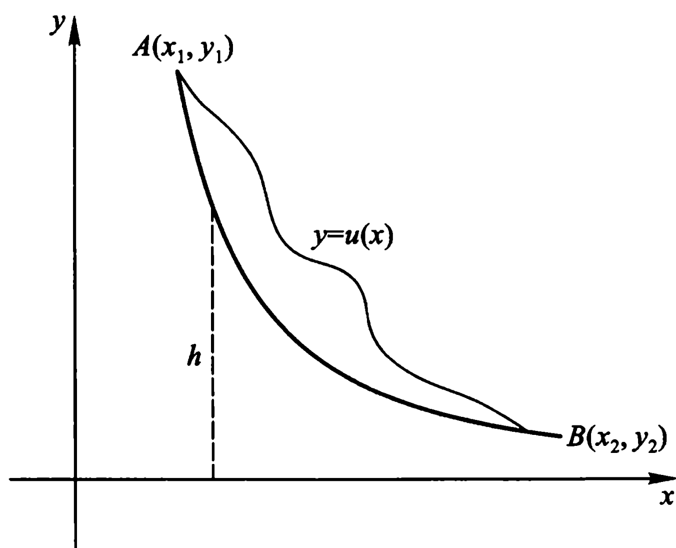


图 1.1 最速下降线

设 $u \in C^1[x_1, x_2], \{(x, u(x)) \mid x \in [x_1, x_2], u(x_i) = y_i, i = 1, 2\}$ 是连接 A, B 的一条曲线. 因为有

$$\begin{cases} \frac{1}{2}mv^2 = mgh, \\ v = \frac{ds}{dt}, \end{cases}$$

所以

$$v = \sqrt{2g(y_1 - u(x))},$$

以及

$$dt = \frac{ds}{v} = \sqrt{\frac{1 + |u'(x)|^2}{2g(y_1 - u(x))}}dx,$$

总时间

$$T = \int_{x_0}^{x_1} dt = \frac{1}{\sqrt{2g}} \int_{x_0}^{x_1} \sqrt{\frac{1 + |u'(x)|^2}{y_1 - u(x)}}dx.$$

令

$$M = \{u \in C^1([x_1, x_2]) \,|\, u(x_i) = y_i, i = 1, 2\},$$

则映射

$$M \to \mathbb{R}, \quad u \mapsto T$$

是一个 "泛函". 这里 u 是自变量, $T = T(u)$ 是因变量. 我们要在 M 中求 u 以使 T 达到极小. □

例 1.2 (测地线) 在单位球面 $S^2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid \sum_{i=1}^3 x_i^2 = 1\}$ 上给定两点 $P_0 = (x_1^0, x_2^0, x_3^0), P_1 = (x_1^1, x_2^1, x_3^1)$, 求连接这两点的弧长最短的曲线.

我们采用球面坐标 $v=(\theta,\varphi)\in[-\frac{1}{2}\pi,\frac{1}{2}\pi]\times[0,2\pi)$, 则

$$\begin{cases} x_1=x_1(v)=\cos\theta\cos\varphi,\\ x_2=x_2(v)=\cos\theta\sin\varphi,\\ x_3=x_3(v)=\sin\theta, \end{cases}$$

确定出 $v^i=(\theta^i,\varphi^i)$, 满足 $P_i=(x_1(v^i),x_2(v^i),x_3(v^i)),i=0,1$.

设 $M=\{v\in C^1([0,1],[-\frac{\pi}{2},\frac{\pi}{2}]\times[0,2\pi))\,|\,v(i)=v^i,i=0,1\}$, 则 $\forall v\in M$, $u(t)=(x_1(v(t)),x_2(v(t)),x_3(v(t)))$ $(t\in[0,1])$ 是连接 P_0,P_1 两点的曲线.

这曲线弧元的平方是

$$ds^2=dx_1^2+dx_2^2+dx_3^2=d\theta^2+\cos^2\theta d\varphi^2=(\theta'(t)^2+\cos^2\theta(t)\varphi'(t)^2)dt^2.$$

故弧长 $L:M\to\mathbb{R}$ 为

$$L(v)=\int_0^1|ds|=\int_0^1\sqrt{(\theta'(t)^2+\cos^2\theta(t)\varphi'(t)^2)}dt.$$

弧长 L 是曲线函数 $v(t)=(\theta(t),\varphi(t))$ 的泛函. 我们要在 M 中求出一个函数 $v(t)=(\theta(t),\varphi(t))$ 使得泛函 $L(v)$ 达到极小值. □

例 1.3 (极小曲面) 在空间 $\mathbb{R}^3$ 中给定一条 Jordan 曲线 Γ, 能否找到一个盘状的曲面 S 张在 Γ 上使其面积达到极小值? 用参数方程描写 $S:(u,v)\mapsto Z=(x,y,z):\bar{D}\to\mathbb{R}^3$, 其中 $D\subset\mathbb{R}^2$ 是单位圆, $u^2+v^2\leqslant 1$,

$$\begin{cases} x=x(u,v),\\ y=y(u,v),\\ z=z(u,v). \end{cases}$$

这曲面的面积是

$$\begin{aligned} A(Z)&=\int_D|Z_u\times Z_v|dudv\\ &=\int_D\sqrt{(x_uy_v-y_ux_v)^2+(y_uz_v-z_uy_v)^2+(z_ux_v-x_uz_v)^2}dudv. \end{aligned}$$

面积 A 是曲面函数 Z 的泛函. 我们要在 $Z|_{\partial D}$ 与 Γ 同胚的条件下求 A 的极小值, 即取

$$M=\{Z\in C^1(\bar{D},\mathbb{R}^3)\,|\,Z|_{\partial D}\simeq\Gamma\},$$

求向量函数 $Z(u,v)=(x(u,v),y(u,v),z(u,v))\in M$, 使得 $A(Z)$ 达到极小值. □

例 1.4 (特征值问题与不等式) 给定 $\mathbb{R}^n$ 中的一个有界区域 Ω. $\forall u \in H_0^1(D)$, 其中 $H_0^1(D)$ 是边值为零的 Sobolev 空间 (详见第十讲), 我们定义能量

$$E(u) = \int_\Omega |\nabla u(x)|^2 dx,$$

以及约束

$$G(u) = \int_\Omega |u(x)|^2 dx = 1.$$

定义

$$M = \{u \in H_0^1(D) \mid G(u) = 1\}.$$

我们要在 M 中寻求函数 u_1 使得泛函 $E: M \to \mathbb{R}$ 在 u_1 达到极小值.

$$\lambda_1 = \min\{E(u) \mid G(u) = 1\}$$

称为第一特征值, u_1 称为第一特征函数. 它们在几何、物理以及工程中都有重要的意义.

许多分析与几何上的不等式都可以提成变分问题求解. 例如, Sobolev 不等式

$$\int_{\mathbb{R}^N} |\nabla u(x)|^2 dx \leqslant S_N \left(\int_{\mathbb{R}^N} |u(x)|^{\frac{2N}{N-2}} dx\right)^{\frac{N-2}{N}},$$

其中

$$S_N = N(N-2)\pi \left(\frac{\Gamma(N)}{\Gamma(N/2)}\right)^{-2/N}$$

我们可以把它化成在

$$M = \left\{u \in H_0^1(\mathbb{R}^N) \,\middle|\, \int_{\mathbb{R}^N} |u(x)|^{\frac{2N}{N-2}} dx = 1\right\}$$

中求泛函

$$I(u) = \int_{\mathbb{R}^N} |\nabla u(x)|^2 dx$$

的极大值.

如其极大值是 S_N, 还表明 S_N 是使这个不等式成立的最佳常数. □

例 1.5 (薄板的振动) 弹性 (齐次、各向同性的) 薄板受外力振动. (所谓薄板是指板的厚度 h 与其最小跨度 a 之比 $h/a \ll 1$. 对于这类弹性力学问题, Kirchoff 提出 "直法线假设", 即薄板在变形时, 它的法线仍保持为法线, 且没有伸长应变.)

设一块薄板占有平面区域 Ω, 它的密度是 $\rho(x, y)$. 用 $w(x, y)$ 表示在 $(x, y) \in \Omega$ 处的位移.

把内力与外力合在一起, 其总效果通过位能密度表现出来.

由应力及应变关系产生的部分位能密度依赖于 $w(x,y)$ 的 Hesse 矩阵

$$\begin{pmatrix} w_{xx} & w_{xy} \\ w_{yx} & w_{yy} \end{pmatrix}.$$

任何物理量都与坐标的选取无关, 所以这部分位能密度只依赖于 $w_{xx}+w_{yy}$ 与 $w_{xx}w_{yy}-w_{xy}^2$ 这两个量.

作用在 Ω 上的外力面密度 $f(x,y)$ 产生的部分位能密度表现为外力功. 如果略去作用在边界 $\partial\Omega$ 上的力与边界弯曲矩不计, 那么总位能

$$U(w)=\int_\Omega \left\{\frac{1}{2}[(w_{xx}+w_{yy})^2-2(1-\mu)(w_{xx}w_{yy}-w_{xy}^2)]+f(x,y)w(x,y)\right\}dxdy,$$

其中 μ 是由材料本身决定的, 称为 Poisson 比.

若我们固定薄板的边界, 即 $w|_{\partial\Omega}=\varphi(x)$, 其中 φ 是 $\partial\Omega$ 上给定的一个函数.

引入 $M=\{w\in C^2(\bar{\Omega})\,|\,w|_{\partial\Omega}=\varphi(x)\}$, 则总位能 U 便是 M 上位移函数 w 的泛函.

力学原理告诉我们: 薄板的平衡 w 遵从变分原理. 也就是说, 那些实现总位能泛函 U 极小的位移函数 w 是薄板的平衡位置. □

§1.4 进一步的例子

除了典型的例子外, 还有一些问题, 从表面上看似乎与泛函极值联系不起来, 但经过一定的转化, 仍可归为变分问题.

例 1.6 (商品的再投资) 生产一种商品, 单位时间产出率为 $q=q(t)$, 生产增长的速度 $\dot{q}$ 与产品的再投资的百分比 $u(t)$ 成正比, 即

$$\dot{q}=\alpha uq,$$

其中 $\alpha>0$ 是一个常数. 因此在时间区间 $[0,T]$ 内市场的总商品量是

$$J(u,q)=\int_0^T(1-u(t))q(t)dt.$$

已知 $q(0)=q_0$, 怎样选择再投资的百分比 $u(t)$, 以使市场的总商品量达到极大值?

在这个问题中, 取

$$M=\{(u,q)\in C[0,T]\times C^1[0,T]\,|\,0\leqslant u(t)\leqslant 1,\dot{q}=\alpha uq,q(0)=q_0\}.$$

$$(u,q) \mapsto J(u,q)$$

是 M 上的一个泛函. 我们要在 M 中求函数 (u,q) 使得泛函 J 达到极大值.

这是一个最优控制问题. 其中 u 和 q 的地位是不平等的, 我们把 u 称为控制变量, q 称为状态变量, 而 J 称为目标泛函. □

我们说在变分问题中自变量是 "函数", 但对于 "函数" 的理解可以相当广泛.

例 1.7 (图像分割) 我们想要从一张 (可能有污损的) 照片中查明人像的边缘. 设此图形占有平面区域 $\Omega \subset \mathbb{R}^2$, 用函数 $g : \Omega \to \mathbb{R}^1$ 表示这图形. 我们要寻求另一个函数 $u : \Omega \to \mathbb{R}^1$, 使之在人像的边缘处与原图像尽量吻合, 而在其余处尽量不要有特殊的标志.

为描写像的边缘, 引入具有有限一维 Hausdorff 测度的闭子集合 $K \subset \bar{\Omega}, \mathrm{H}^1(K) < \infty$, 其中 $\mathrm{H}^1(K)$ 是 K 的一维 Hausdorff 测度. 定义

$$I(K,u) = \int_{\Omega\backslash K} |\nabla u|^2 dxdy + \mu \int_{\Omega\backslash K} |u-g|^2 dxdy + \lambda \mathrm{H}^1(K),$$

其中 $\lambda, \mu > 0$ 是可以调节的参数.

值得注意的是: 这里 I 不仅依赖于函数 u, 而且还依赖于闭子集 K. 虽然这个未知变量 K 是一个闭集合, 不是一个函数, 但是我们可以用它的特征函数 χ_K 来替代:

$$\chi_K(x) = \begin{cases} 1, & x \in K, \\ 0, & x \in \Omega\backslash K. \end{cases}$$

I 照样可以看成是依赖于 χ_K 与 u 的一个泛函. 今后在不会发生混淆的情况下, 我们不再区别 K 与 χ_K.

我们可以把图像分割问题提成求泛函 I 的极值问题. 为此定义自变量的集合

$$M = \left\{ (K,u) \,\middle|\, K \subset \Omega \text{ 是闭子集}, \mathrm{H}^1(K) < \infty, \int_\Omega (|\nabla u|^2 + |u|^2) dxdy < \infty \right\},$$

其中 ∇u 按广义函数的意义来理解.

我们要在 M 中寻求 (K,u), 使得

$$I : M \to \mathbb{R}$$

达到极小值. □

例 1.8 (调和映射) 设 (M,g), (N,h) 是两个紧 Riemann 流形. 给定一个映射 $u \in C^1(M,N)$, 其微分 $du \in \Gamma(T^*M \times u^{-1}TN)$ 是乘积向量丛的截面, 其中

$u^{-1}TN$ 是 M 上的向量丛, 度量为 $h\circ u$, 而 T^*M 是 M 的余切丛. 用局部坐标来写: $x=(x^1,\cdots,x^m)$, $u=(u^1,\cdots,u^n)$, 有

$$du=\sum_{i=1}^{n}\sum_{\alpha=1}^{m}\frac{\partial u^i}{\partial x^\alpha}dx^\alpha\otimes\frac{\partial}{\partial u^i}.$$

如果我们定义能量密度为

$$L(x,u,du)=\frac{1}{2}|du(x)|_h^2=\frac{1}{2}\sum_{i,j=1}^{n}\sum_{\alpha,\beta=1}^{m}h_{ij}(u(x))g^{\alpha\beta}(x)\frac{\partial u^i}{\partial x^\alpha}\frac{\partial u^j}{\partial x^\beta},$$

dV_g 为 M 上的体积元, 那么能量

$$E(u)=\int_M L(x,u(x),du(x))dV_g$$

是映射 u 的一个泛函.

映射 u 本身也可以看成是一种 "函数". 取 $\mathcal{M}=C^1(M,N)$, 则 $E:\mathcal{M}\to\mathbb{R}$. 使得泛函 E 在 $\mathcal{M}$ 中达到极小值的映射, 称为调和映射. 它在微分几何学中有重要的意义. □

变分问题不一定只研究泛函的极大值或极小值, 也可以研究泛函的其他临界值.

例 1.9 (Hamilton 方程组) 给定一个函数 $H\in C^1(\mathbb{R}^1\times\mathbb{R}^n\times\mathbb{R}^n)$. 用以下记号: $x=(x_1,\cdots,x_n)$, $p=(p_1,\cdots,p_n)$, $(t,x,p)\in\mathbb{R}^1\times\mathbb{R}^n\times\mathbb{R}^n$, 以及 $(x,p)_{\mathbb{R}^n}=\sum_{i=1}^n x_ip_i$.

当 x,p 都是 t 的 (向量值) 函数时, 常微分方程组

$$\begin{cases}\dot{x}=H_p(t,x,p),\\ \dot{p}=-H_x(t,x,p)\end{cases}$$

称为 Hamilton 方程组.

引入 $\mathcal{M}$ 为 $C^1(\mathbb{R}^1,\mathbb{R}^n\times\mathbb{R}^n)$ 中函数连同其导数都是 $\mathbb{R}^1$ 上可积的函数类, 并且对于函数 H 添加适当限制, 使得

$$I(x,p)=\int_{\mathbb{R}^1}[(\dot{x},p)_{\mathbb{R}^n}-H(t,x,p)]dt$$

是 $\mathcal{M}$ 上定义的泛函.

显然 I 既无上界又无下界, 所以既没有极大值也没有极小值. 但在下几讲中, 我们将看到: 对于适当的 $\mathcal{M}$, 使得 I 达到临界值 (又称为稳定值) 的 (向量值) 函数 $(x(t),p(t))$ 正是 Hamilton 方程组的解. □

例 1.10 (Einstein 方程) 在广义相对论中, 用空间 $\mathbb{R}^4$ 的一个 (1,3) 型 Minkowski 度量 (g_{ij}) 来描写重力场. 这个 Minkowski 度量是取值于符号为 $(1,-1,-1,-1)$ 的对称矩阵函数. 对应的线元是

$$ds^2 = \sum_{i,j} g_{ij}dx^i dx^j.$$

设欧几里得空间坐标是 (x,y,z), 时间坐标是 t, 则 $(x^0,x^1,x^2,x^3)=(ct,x,y,z)$. 在 "惯性标架" 下, 我们有

$$ds^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2.$$

用 Γ^i_{jk} 表示与此度量相关的联络. 这时我们有曲率张量

$$R_{iklm} = \frac{1}{2}\left(\frac{\partial^2 g_{im}}{\partial x^k \partial x^l} + \frac{\partial^2 g_{kl}}{\partial x^i \partial x^m} - \frac{\partial^2 g_{il}}{\partial x^k \partial x^m} - \frac{\partial^2 g_{km}}{\partial x^i \partial x^l}\right) + \sum_{n,p} g_{np}(\Gamma^n_{kl}\Gamma^p_{im} - \Gamma^n_{km}\Gamma^p_{il}),$$

对应的 Ricci 张量是

$$R_{ik} = \sum_l \left(\frac{\partial \Gamma^l_{ik}}{\partial x^l} - \frac{\partial \Gamma^l_{il}}{\partial x^k}\right) + \sum_{l,m}\left(\Gamma^l_{ik}\Gamma^m_{lm} - \Gamma^m_{il}\Gamma^l_{km}\right),$$

数量曲率是

$$R = \sum_{i,k} g^{ik}R_{ik} = \sum_{i,k}\sum_{l,m} g^{il}g^{km}R_{iklm}.$$

Einstein 引入的作用量 $S=S(g)$ 由两部分组成:

$$S = S_g + S_m,$$

其中

$$S_g = \int R d\Omega, \quad d\Omega = \sqrt{|\det(g_{ij})|}d^4x$$

称为场的 Hilbert 作用量场, 它表示没有物质时场的贡献. 而

$$S_m = \frac{1}{c}\int \Lambda\sqrt{-\det(g_{ij})}d^4x,$$

其中 Λ 是由物质与度量决定的一个函数, 它表示在场中物质的贡献.

在这里, 度量 g 是自变量, 它是非退化对称 (1,3) 型的矩阵值函数, 作用量 $S=S(g)$ 是 g 的泛函.

物质的运动遵从变分原理, 即重力场 —— Minkowski 度量 —— 使得作用量 S 达到稳定. □

总之, 变分问题内容非常丰富. 从经典力学到规范场论, 物质运动的规律遵从变分原理. 例如, 光线的折射, 毛细管中液质的弯曲, 以及肥皂薄膜张成的曲面都是我们日常看到的遵从变分原理的自然现象. 此外, 从工程设计到社会经济生活, 人们在一定条件下追求最快的速度、最大的射程、最低的损耗、最优的形状、最轻的载体、最高的效益、最清晰的图像等都涉及泛函的极值问题.

第二讲　Euler-Lagrange 方程

§2.1　函数极值必要条件之回顾

在研究泛函极值的必要条件之前, 先回顾一下函数极值的必要条件. 设 $\Omega \subset \mathbb{R}^N$ 是一个开集, 又设函数 $f \in C^1(\Omega)$ 在 $x_0 \in \Omega$ 达到极小值, 即有 x_0 的一个开邻域 $U \subset \Omega$ 使得

$$f(x) \geqslant f(x_0), \quad \forall\, x \in U.$$

于是 $\forall\, h \in \mathbb{R}^n \backslash \{\theta\}$, 其中 θ 是零向量, $\exists\, \varepsilon(h) > 0$, 当 $0 < |\varepsilon| < \varepsilon(h)$ 时, $x_0 + \varepsilon h \in U$, 有

$$f(x_0 + \varepsilon h) \geqslant f(x_0),$$

即

$$\frac{1}{\varepsilon}[f(x_0 + \varepsilon h) - f(x_0)] \geqslant 0.$$

令 $\varepsilon \to 0$,

$$(\nabla f(x_0), h)_{\mathbb{R}^n} = 0.$$

而 $h \in \mathbb{R}^n \backslash \{\theta\}$ 是任意的, 所以

$$\nabla f(x_0) = 0.$$

这就是 x_0 成为 f 的极小点的必要条件.

§2.2 Euler-Lagrange 方程的推导

在这一讲, 我们只讨论泛函依赖于单变量 (可以是向量值) 函数的情形. 对于多变量函数的问题留待第六讲去讨论. 现在给定一个区间 $J=[t_0,t_1]\subset\mathbb{R}^1$ 和一个开区域 $\Omega\subset\mathbb{R}^N$. 给定一个连续可微函数 $L=L(x,u,p)$, $L\in C^1(J\times\Omega\times\mathbb{R}^N,\mathbb{R}^1)$, 再给定两个点 $P_0,P_1\in\Omega$. 令

$$M=\{u\in C^1(J,\Omega)\,|\,u(t_i)=P_i, i=0,1\},$$

以及 M 上的泛函

$$I(u)=\int_J L(t,u(t),\dot{u}(t))dt.$$

称 u^* 是 I 在 M 上的极小点, 如果存在 u^* 的一个邻域 U (在这里指的是在 M 中按 $C^1(J,\Omega)$ 拓扑) 使得:

$$I(u)\geqslant I(u^*),\quad \forall u\in M\cap U.$$

我们要在 u^* 存在的前提下, 寻求使泛函 I 在函数 u^* 达到极小值应满足的必要条件.

类似于函数极值问题, $\forall\varphi\in C_0^1(J,\mathbb{R}^N)$, $C_0^1(J,\mathbb{R}^N)$ 表示 $C_0^\infty((t_0,t_1),\mathbb{R}^N)$ 在 $C^1(J,\mathbb{R}^N)$ 中的闭包, $\exists\varepsilon=\varepsilon(\varphi)>0$ 使得当 $0<|\varepsilon|<\varepsilon(\varphi)$ 时 $u^*+\varepsilon\varphi\in U$, 便有

$$I(u^*+\varepsilon\varphi)-I(u^*)\geqslant 0.$$

此即

$$\begin{aligned}
\delta I(u^*,\varphi)&=\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}(I(u^*+\varepsilon\varphi)-I(u^*))\\
&=\int_J\sum_{i=1}^{n}[L_{u^i}(t,u^*(t),\dot{u}^*(t))\varphi^i(t)+L_{p^i}(t,u^*(t),\dot{u}^*(t))\dot{\varphi}^i(t)]dt\\
&=\int_J\left(\sum_{i=1}^{N}\int_{t_0}^{t}L_{u^i}(s,u^*(s),\dot{u}^*(s))ds-L_{p^i}(t,u^*(t),\dot{u}^*(t))\right)\dot{\varphi}^i(t)dt\\
&\geqslant 0,
\end{aligned}$$

$\forall\varphi\in C_0^1(J,\mathbb{R}^N)$. 我们把 $\delta I(u^*,\varphi)$ 称为 I 对 φ 的一阶变分.

将 $\varepsilon>0$ 换成 $\varepsilon<0$ 相当于在此式中将 φ 换成 $-\varphi$, 于是得到等式, 即 $\forall\varphi\in C_0^1(J,\mathbb{R}^N)$,

$$\boxed{\int_J\left(\sum_{i=1}^{N}\int_{t_0}^{t}L_{u^i}(s,u^*(s),\dot{u}^*(s))ds-L_{p^i}(t,u^*(t),\dot{u}^*(t))\right)\dot{\varphi}^i(t)dt=0.}$$

剩下来我们要把这个积分等式中的任意函数 φ 去掉, 得到一个关于 u^* 的关系式. 为此需要下述引理.

引理 2.1 (du Bois-Reymond) 若 $\psi \in C[t_0, t_1]$, 且

$$\int_J \psi(t) \cdot \dot{\lambda}(t) dt = 0, \quad \forall \lambda \in C_0^1(J),$$

其中 $C_0^1(J) = \{u \in C^1(J) \mid u(t_0) = u(t_1) = 0\}$, 则 $\psi = \text{const}$.

证明 令 $c = \frac{1}{|J|} \int_J \psi(t) dt$ 且 $\lambda(t) = \int_{t_0}^t (\psi(s) - c) ds$, 则 $\lambda \in C_0^1(J)$. 因此,

$$\int_J (\psi(t) - c)^2 dt = \int_J \psi(t)(\psi(t) - c) dt = \int_J \psi(t) \cdot \dot{\lambda}(t) dt = 0.$$

因为 ψ 是连续的, 所以 $\psi = \text{const}$. □

于是我们得到

定理 2.1 设 $u^* \in M$ 是泛函 I 在 M 上的一个极小点, 则它满足下列积分形式的 Euler-Lagrange 方程(简称 E-L 方程):

$$\boxed{\int_{t_0}^t L_{u^i}(s, u(s), \dot{u}(s)) ds - L_{p^i}(t, u(t), \dot{u}(t)) = \text{const}, \quad 1 \leqslant i \leqslant N, \quad \forall t.}$$

E-L 方程是泛函极小值的必要条件, 显然它不是充分的.

E-L 方程的解是相应泛函 I 的临界点, 它可以是极大值或极小值, 也可以是其他形式的临界点.

注 2.1 当 $L \in C^1, u \in C^1$ 时, 在广义函数意义下, 积分形式的 E-L 方程又可以改写成微分形式的 E-L 方程:

$$\boxed{-DL_p(t, u(t), \dot{u}(t)) + L_u(t, u(t), \dot{u}(t)) = 0.}$$

其中 D 是广义导数. 我们定义 Euler-Lagrange 算子E_L 如下:

$$(E_L u)(t) = DL_p(t, u(t), \dot{u}(t)) - L_u(t, u(t), \dot{u}(t)).$$

特别地, 如果 $L \in C^2$ 以及 $u \in C^2$, 那么上式在逐点意义下成立, 也就是说可以把 D 换为普通导数 $\frac{d}{dt}$. □

注 2.2 我们可以把定理 1.1 中的函数类 $C^1(J)$ 放大一些. 例如, 考虑 Lipschitz 函数类 $\mathrm{Lip}(J)$. 因为 Lipschitz 函数是绝对连续的, 所以几乎处处存在导数 $\dot{u}(t)$, 而且泛函 I 中的积分可以按 Lebesgue 积分意义理解. 取

$$M = \{u \in \mathrm{Lip}(J, \Omega) \,|\, u(t_i) = P_i, i = 0, 1\}.$$

注意: 在 M 中, $\forall \delta > 0, U = \{v \in \mathrm{Lip}(J) \,|\, |v(t) - u^*(t)| + |\dot{v}(t) - \dot{u}^*(t)| < \delta, \text{a.e.}\, t \in J\}$ 是 u^* 的一个邻域.

E-L 方程仍然在 Lebesgue 意义下对几乎所有的 $t \in J$ 成立:

$$\int_{t_0}^{t} L_{u^i}(s, u(s), \dot{u}(s))ds - L_{p^i}(t, u(t), \dot{u}(t)) = \mathrm{const}, \quad 1 \leqslant i \leqslant n.$$

事实上, 因为这时 $\dot{u}$ 几乎处处存在且有界, 从而存在 $(u(t), \dot{u}(t))$ 的一个紧邻域 $W \subset \Omega \times \mathbb{R}^N$, 使得 L 的导数在 $J \times W$ 上有界. 我们得到

$$\begin{aligned}
\delta I(u^*, \varphi) &= \lim_{s \to} \frac{1}{s}(I(u^* + s\varphi) - I(u^*)) \\
&= \lim_{s \to} \frac{1}{s} \int_J [(L(t, u^*(t) + s\varphi(t), \dot{u}^*(x) + s\dot{\varphi}(x)) - L(t, u^*(t), \dot{u}^*(t))]dt \\
&= \int_J \sum_{i=1}^{n} [L_{u^i}(t, u^*(t), \dot{u}^*(t))\varphi^i(t) + L_{p^i}(t, u^*(t), \dot{u}^*(t))\dot{\varphi}^i(t)]dt.
\end{aligned}$$

这是因为积分号内的差有一致的控制. 求导过程可以用 Lebesgue 控制定理来保证积分号下取极限.

此外, 在 du Bois-Reymond 引理中我们可以把 $\psi \in C^1(J)$ 换成 $\psi \in L^\infty(J)$, 把 $\lambda \in C_0^1(J)$ 换成 $\lambda \in AC_0(J)$, 即边值为 0 的 J 上的绝对连续函数空间. □

注 2.3 逐段 C^1 的连续函数 (即存在有限点集 $D = \{a_1, \cdots, a_k\}$ 使得 $u \in C^1(J \backslash D)$, 而且 $\dot{u}(a_i \pm 0)$ 都存在, $i = 1, \cdots, k$) 是 Lipschitz 函数. 因此积分形式的 E-L 方程对于逐段 C^1 的连续函数类成立. □

力学、几何与物理中的基本方程大都是变分问题的 E-L 方程, 例如:

例 2.1 (质点运动方程) 受外力 F 作用的质量为 m 的质点, 设其位置坐标为 $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, 则速度 $v = \dot{x}$, 动能 $T = \frac{1}{2}mv^2$. 倘若外力 F 有位势, 即存在函数 V 满足

$$-\nabla V = F,$$

我们称

$$L = T - V = \frac{1}{2}mv^2 - V = \frac{1}{2}m|\dot{x}|^2 - V(x)$$

为 Lagrange 函数 (Lagrangian). 适当确定定义域 M, 考虑泛函

$$I(x)=\int_{t_1}^{t_2}L(x(t),\dot{x}(t))dt,$$

使得 I 实现极小的轨道 $x(t)$ 满足 Euler-Lagrange 方程

$$F=m\ddot{x}.$$

这正是 Newton 第二定律所确定的运动轨道.

对于有 n 个自由度的质点组, 位置坐标记作 $q=(q_1,\cdots,q_n)$,

- 动能 $T=\frac{1}{2}\sum_{i,j=1}^{n}a_{ij}\dot{q}_i\dot{q}_j$, 其中 (a_{ij}) 是正定阵,
- 位能 $V=V(q_1,\cdots,q_n)$,
- Lagrange 函数 $L=T-V$,
- 泛函

$$I(q)=\int_{t_0}^{t_1}L(q(t),\dot{q}(t))dt=\int_{t_0}^{t_1}\left(\frac{1}{2}\sum_{i,j=1}^{n}a_{ij}\dot{q}_i\dot{q}_j-V(q_1,\cdots,q_n)\right)dt,$$

由此导出 E-L 方程组为

$$\sum_{j=1}^{n}a_{ij}\ddot{q}_j=-\frac{\partial V}{\partial q_i}\quad(i=1,2,\cdots,n).$$

这还是 Newton 方程. □

例 2.2 (测地线) 设 (M,g) 是一个 N 维 Riemann 流形, Riemann 度量为 $(g_{ik}(u))$, 其中 g_{ik} 是一个 $N\times N$ 正定对称矩阵. 当 P_1,P_2 两点在 M 的同一个坐标卡 $U\subset M$ 内时, U 同胚于 $\mathbb{R}^N$ 中的一个开集. 令 $L:\mathbb{R}^n\times\mathbb{R}^N\to\mathbb{R}^1$,

$$L(u,p)=\sum_{i,j=1}^{N}g_{ij}(u)p_ip_j.$$

对于泛函

$$I(u)=\int_J L(u(x),\dot{u}(x))dx,$$

有 E-L 方程

$$\frac{d}{dt}\sum_j\{g_{ij}(u)\dot{u}^j\}=\frac{1}{2}\sum_{j,k}g_{jk,i}(u)\dot{u}^j\dot{u}^k,$$

其中

$$g_{jk,i}(u) = \frac{\partial}{\partial u^i} g_{jk}(u).$$

所以

$$\sum_j g_{ij}(u)\ddot{u}^j + \sum_{j,k} \left(g_{ij,k}(u)\dot{u}^j\dot{u}^k - \frac{1}{2} g_{kj,i}\dot{u}^j\dot{u}^k \right) = 0.$$

注意到

$$\sum_{j,k} g_{lj,k}(u)\dot{u}^k\dot{u}^j = \sum_j \frac{d}{dt} g_{lj}(u)\dot{u}^j = \sum_k \frac{d}{dt} g_{lk}\dot{u}^k = \sum_{j,k} g_{kl,j}(u)\dot{u}^k\dot{u}^j,$$

我们引入 Christofel 符号,

$$\Gamma_{jlk}(u) = \frac{1}{2}\{g_{lj,k} + g_{kl,j} - g_{jk,l}\}.$$

便有

$$\sum_j g_{ij}(u)\ddot{u}^j + \sum_{jk} \Gamma_{jik}(u)\dot{u}^k\dot{u}^j = 0.$$

因为 (g_{ik}) 是可逆的, 记其逆阵为 (g^{ik}), 又记

$$\Gamma^i_{jk} = \sum_l g^{il}\Gamma_{jlk}.$$

于是 E-L 方程又可以写成:

$$\ddot{u}^i + \sum_{j,k} \Gamma^i_{jk}(u)\dot{u}^j\dot{u}^k = 0, \quad \forall\, i.$$

这就是微分几何中的测地线方程. □

在后面几讲中, 我们还将利用 E-L 方程导出许多物理基本方程.

- **变分导数**

以上 E-L 方程的推导过程虽然是在整个区间 J 上进行的, 但 $\forall c \in \mathrm{int}(J)$, 在这点的 E-L 方程实际上只依赖于这点邻近 L 的行为, 即, 只依赖于包含 c 的任意邻域 $(c-h, c+h) \subset \mathrm{int}(J)$. 因为支集在 $(c-h, c+h)$ 内的试探函数 φ 显然在 $C_0^1(J, \mathbb{R}^N)$ 内, 由这些函数的任意性就导出了在 $(c-h, c+h)$ 上的 E-L 方程 (见图 2.1).

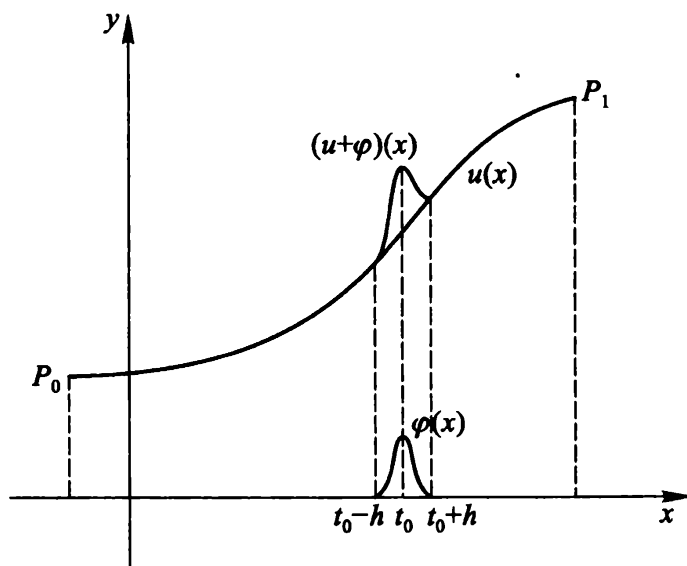


图 2.1 Euler-Lagrange 方程

我们还可以通过以下极限过程来看这种局部性. 以 $N=1$ 为例, 并假设 $L\in C^2$, $u\in C^2$.

$$\begin{aligned}\lim\frac{I(u+\varphi)-I(u)}{\Delta\sigma}&=-\lim\frac{\displaystyle\int_{c-h}^{c+h}\int_{t_0}^{t}E_L(u+\theta\varphi)(s)ds\dot{\varphi}(t)dt}{\Delta\sigma}\\&=\lim\frac{\displaystyle\int_{c-h}^{c+h}E_L(u+\theta\varphi)(t)\varphi(t)dt}{\Delta\sigma}\\&=E_L(u)(c),\end{aligned}$$

其中 $\theta=\theta(t)\in(0,1)$, φ 的支集在 $(c-h,c+h)$ 内, $\Delta\sigma=\int_{c-h}^{c+h}\varphi(t)dt$ 是曲线 $u(t)+\varphi(t)$ 与曲线 $u(t)$ 在 $t\in(c-h,c+h)$ 之间所夹的面积; 而极限过程是

$$h\to 0,\quad \sup_{t\in[c-h,c+h]}|\dot{\varphi}(t)|\to 0.$$

在这个意义上, 我们称 Euler-Lagrange 算子对函数 u 作用后在 t 点的值

$$E_L(u)(t)=\frac{d}{dt}L_p(t,u(t),\dot{u}(t))-L_u(t,u(t),\dot{u}(t))$$

为 I 在 t 的变分导数.

§2.3 边值条件

在 §2.1 中, 定义域

$$M=\{u\in C^1(J,\Omega)\,|\,u(t_i)=P_i,i=0,1\}$$

中的函数 u 在 J 的两个端点 t_0, t_1 取定值 $u(t_i) = P_i, i = 0, 1$.

如此得到的 u^* 除了满足 E-L 微分方程

$$-\frac{d}{dt}L_p(t, u(t), \dot{u}(t)) + L_u(t, u(t), \dot{u}(t)) = 0$$

外, 还必须满足边值条件

$$u(t_0) = P_0, \quad u(t_1) = P_1.$$

如果把定义域换成 $M = C^1(J, \mathbb{R}^N)$, 也就是说, 在区间 J 的两个端点不加任何限制, 那么泛函 I 的极值函数 u^* 应该满足怎样的方程和边值条件呢?

回想在 §2.1 推导 E-L 方程的过程, 关键在于选取 u^* 邻近的函数 u 来与 u^* 在泛函值上作比较. 那时我们取

$$u = u^* + \varepsilon\varphi,$$

其中 $\varphi \in C_0^1(J, \mathbb{R}^N)$, 因为 φ 在端点为零, 所以 u 保持与 u^* 有同样的端点值.

对于现在这种情形, 没有必要对 φ 的端点值作任何限制, 也就是说所有 $C^1(J, \mathbb{R}^N)$ 中的函数 φ 都可以用来构造 u. 倘若 $u^* \in C^2(J, \mathbb{R}^N)$, 那么由分部积分公式, 我们有

$$\begin{aligned}\delta I(u^*, \varphi) &= \int_J \left(\sum_{i=1}^N \int_{t_0}^t L_{u^i}(s, u^*(s), \dot{u}^*(s))ds - L_{p^i}(t, u^*(t), \dot{u}^*(t)) \right) \dot{\varphi}^i(t)dt \\ &= -\int_J \left(\sum_{i=1}^N [L_{u^i}(t, u^*(t), \dot{u^*}(t)) - \frac{d}{dt}L_{p^i}(t, u^*(t), \dot{u}^*(t))] \right) \varphi^i(t)dt \\ &\quad - \sum_{i=1}^N [L_{p^i}(t, u^*(t), \dot{u}^*(t))\varphi^i(t)|_{t_0}^{t_1}].\end{aligned}$$

注意到 $C_0^1(J, \mathbb{R}^N) \subset C^1(J, \mathbb{R}^N)$, 所以在上式中先取任意的 $\varphi \in C_0^1(J, \mathbb{R}^N)$, 得到同样的 E-L 方程, 再任意选取 $\varphi \in C^1(J, \mathbb{R}^N)$ (实际上只是任意选取 φ 在端点的值). 因为此时右端第一项已经消失, 而第二项中 $\varphi^i(t_j)$ $(j = 0, 1, i = 1, \cdots, N)$ 是任意的, 所以有

$$L_{p^i}(t_j, u^*(t_j), \dot{u}^*(t_j)) = 0, \quad i = 1, \cdots, N, \ j = 0, 1.$$

当然, M 还可以有许多种不同的选取, 例如一个端点的函数值固定, 另一端不固定; 也可以对向量值函数的不同分量采用不同的端点条件.

还有一些其他类型的边值条件, 如周期条件、自由边值等, 在后面几讲再专门讨论.

注意: 在以上边值条件的讨论中, 我们始终假设定义域 M 中的函数都是连续可微的. 如果把连续可微函数类换成逐段连续 C^1 的函数类, 虽然 E-L 方程不变 (局部性), 但在导数的跳跃点上需要添加角点条件 (corner condition) (见习题 2, 4).

§2.4 求解 Euler-Lagrange 方程的例子

当 $N=1$ 时, 在以下几种特殊情形 E-L 方程可以简化.

- **情形 1.** L 不含 u, $L=L(t,p)$.

这时
$$\frac{d}{dt}L_p(t,\dot{u}(t))=0.$$

$L_p(t,\dot{u}(t))=c$ 是一个不含 u 的一阶方程, 如果能解出 $\dot{u}$ (例如 $L_{pp}(t,p)\neq 0$), 得到
$$\dot{u}(t)=g(t,c),$$
那么也就可以积分出 $u(t)$.

例 2.3 设 $M=\{u\in C^1([1,2])\,|\,u(1)=0,u(2)=1\}$,
$$I(u)=\int_1^2\sqrt{1+\dot{u}^2}\frac{dt}{t}.$$
求使泛函 I 达到极小的函数 u. □

解 因为 $L=t^{-1}\sqrt{1+p^2}$, 所以
$$L_p=\frac{p}{t\sqrt{1+p^2}}=C.$$
即得
$$\dot{u}^2(1-C^2t^2)=C^2t^2 \quad 或 \quad \dot{u}=\frac{Ct}{\pm\sqrt{1-C^2t^2}},$$
把符号吸收到常数 C 中去, 再积分
$$u=\frac{1}{C}\sqrt{1-C^2t^2}+C_1.$$
代入边值条件, 得到 $C=\frac{1}{\sqrt{5}},C_1=2$. 求得解
$$(u-2)^2+t^2=5.$$

• **情形 2 (自守系统).** L 与 t 无关, $L = L(u,p)$.

引入 Hamilton 量

$$H(u,p) = pL_p(u,p) - L(u,p).$$

定理 2.2 设 $L \in C^2$, 且与 t 无关. 又设 $u \in C^2(J,\mathbb{R}^1)$ 是 E-L 方程的解, 则我们有

$$H(u(t),u'(t)) = \text{const}, \quad \forall t.$$

证明 直接计算

$$\frac{d}{dt}H(u(t),u'(t)) = u'(t) \cdot E_L(u)(t) = 0.$$

例 2.4 (最速下降线) 这时,

$$L(u,p) = \frac{1}{\sqrt{2g}}\frac{\sqrt{1+p^2}}{\sqrt{y_1-u}}.$$

利用定理 2.2, 有

$$pL_p - L|_{(u,u')} = \text{const}.$$

即有常数 c 使得

$$-\frac{\sqrt{1+u'^2}}{\sqrt{y_1-u}} + \frac{u'^2}{\sqrt{(1+u'^2)(y_1-u)}} = c.$$

由此推出

$$c^2(1+u'^2)(y_1-u) = 1.$$

令 k 为一待定正常数, 作变量替换, 引入参变量 θ,

$$\begin{cases} x = x(\theta), \\ u = u(\theta), \end{cases}$$

并令

$$u(\theta) = y_1 - k(1-\cos\theta),$$

则

$$c^2\left(1 + k^2\frac{\sin^2\theta}{\dot{x}(\theta)^2}\right)k(1-\cos\theta) = 1.$$

取 $c = \sqrt{\frac{1}{2k}}$, 有

$$\dot{x}(\theta) = k(1-\cos\theta).$$

我们得到

$$\begin{cases} x(\theta) = x_1 + k(\theta - \sin\theta), \\ u(\theta) = y_1 - k(1 - \cos\theta), \\ \theta \in [0, \Theta]. \end{cases}$$

再通过

$$\begin{cases} x(\Theta) = x_2, \\ u(\Theta) = y_2 \end{cases}$$

确定 k 与 Θ. □

例 2.5 (旋转极小曲面) 在平面上给定两点 $P_1 = (x_1, y_1), P_2 = (x_2, y_2)$, $y_1, y_2 > 0, x_1 < x_2$. 求连接这两点的一个函数 $u \in C^1[x_1, x_2]$, 使得它的图像绕 x 轴旋转后, 所得到的旋转曲面的面积最小.

设 $u(x_i) = y_i, i = 1, 2$, 并且 $u(x) > 0$, 其旋转曲面的面积是:

$$I(u) = 2\pi \int_{x_1}^{x_2} u(x)\sqrt{1 + u'(x)^2}dx.$$

我们要求 u 以使 I 达到极小值.

由于 Hamilton 量守恒,

$$L(u(t), p(t)) - \dot{u}(t)L_p(u(t), p(t)) = C,$$

即

$$u(t)\sqrt{1 + \dot{u}^2(t)} - \frac{u(t)\dot{u}^2(t)}{\sqrt{1 + \dot{u}^2(t)}} = C, \quad \forall t,$$

或

$$\dot{u} = C^{-1}\sqrt{u^2 - C^2}.$$

积分后, 得到

$$C \ln \frac{u + \sqrt{u^2 - C^2}}{C} = x + C_1.$$

故有

$$u = C \cosh \frac{x + C_1}{C}.$$

这是悬链线方程. □

例 2.6 (球面上的测地线) 回到例 1.2, 用 φ 作参数. 利用定理 2.2, 将 E-L 方程写成

$$\frac{\theta'^2(\varphi)}{\sqrt{\cos^2\theta(\varphi) + \theta'(\varphi)^2}} - \sqrt{\cos^2\theta(\varphi) + \theta'(\varphi)^2} = c,$$

其中 c 是一个常数, 由定义, $-1 \leqslant c < 1$. 于是

$$-\cos^2\theta(\varphi) = c\sqrt{\cos^2\theta(\varphi) + \theta'(\varphi)^2},$$

即

$$c^2\theta'^2 = \cos^4\theta - c^2\cos^2\theta,$$

用变量替换 $t = \tan\theta$,

$$\pm\varphi + \varphi_0 = \int \frac{cd\theta}{\cos\theta\sqrt{\cos^2\theta - c^2}} = \int \frac{cdt}{\sqrt{(1-c^2) - c^2t^2}} = \arcsin\frac{ct}{\sqrt{1-c^2}},$$

其中 $0 < c^2 < 1$, 所以

$$\tan\theta(\varphi) = \frac{\sqrt{1-c^2}}{c}\sin(\pm\varphi + \varphi_0)$$

或

$$\theta(\varphi) = \arctan\frac{\sqrt{1-c^2}}{c}\sin(\pm\varphi + \varphi_0).$$

这两个待定常数 c, φ_0 由 P_0, P_1 确定.

这是球面上大圆的方程. 当 $c = 0$ 时, $\theta = \pm\pi/2$, 它只对应于南、北两极, 不可能是曲线. 当 $c = -1$ 时, 对应于赤道. □

• **情形 3.** L 不含 p, $L = L(t, u)$.

这时 E-L 方程成为一个函数方程

$$L_u(t, u) = 0.$$

其解是一条或几条曲线.

例 2.7 $I(u) = \int_a^b (t-u)^2 dt$. 其 E-L 方程是

$$t - u = 0,$$

即直线 $u = t$. □

• **坐标变换**

利用变分导数的表达式, 我们来证明 E-L 方程在变量替换下是不变的. 设

$$\begin{cases} s = s(t, u), \\ v = v(t, u), \end{cases}$$

逆变换为

$$\begin{cases} t = t(s, v), \\ u = u(s, v), \end{cases}$$

则 Lagrange 函数 L 变为

$$\tilde{L}(s, v, q) = L\left(t(s, v), u(s, v), \frac{u_s + u_v q}{t_s + t_v q}\right)(t_s + t_v q).$$

若 t 的区间 $[t_0, t_1]$ 变为 s 的区间 $[s_0, s_1]$, 则

$$\int_{t_0}^{t_1} L(t, u(t), \dot{u}(t))dt = \int_{s_0}^{s_1} \tilde{L}(s, v(s), \dot{v}(s))ds,$$

E-L 方程变为

$$\tilde{L}_v - \frac{d}{ds}\tilde{L}_q = 0.$$

可以先求解后一个方程, 再通过前面的变换回到原问题的解.

例 2.8 考虑以下泛函的极值:

$$I(r) = \int_{\varphi_0}^{\varphi_1} \sqrt{r^2 + \dot{r}^2} d\theta,$$

其中 $r = r(\theta)$.

对应的 E-L 方程是

$$\frac{r}{\sqrt{r^2 + \dot{r}^2}} - \frac{d}{d\theta}\frac{\dot{r}}{\sqrt{r^2 + \dot{r}^2}} = 0.$$

作坐标变换

$$x = r\cos\theta, \quad u = r\sin\theta,$$

泛函变为

$$I(u) = \int_{x_0}^{x_1} \sqrt{1 + \dot{u}^2} dx.$$

其 E-L 方程是

$$\ddot{u} = 0.$$

解出通解

$$u = ax + b.$$

代回到原问题得

$$r\sin\theta = ar\cos\theta + b.$$

习 题

1. 给定一个区间 $J=[t_0,t_1]\subset\mathbb{R}^1$ 和一个开区域 $\Omega\subset\mathbb{R}^N$. 给定一个连续可微函数 $L=L(x,u,p)$, $L:\in C^1(J\times\Omega\times\mathbb{R}^N,\mathbb{R}^1)$. 给定两个向量 $\xi_0,\xi_1\in\mathbb{R}^N$, 令

$$M_1=\{u\in C^1(J,\mathbb{R}^N)\,|\,\dot{u}(t_i)=\xi_i,\ i=0,1\}.$$

定义 M_1 上的泛函

$$I(u)=\int_J L(t,u(t),\dot{u}(t))dt.$$

设 $u_0\in M_1$ 使 I 达到极小, 问: u_0 应满足的必要条件是什么?

2. (第一 Erdmann 角点条件) 在习题 1 假设的条件下, 设 $P_0,P_1\in\Omega$,

$$u_0\in M_2=\{u\in PWC^1(J,\mathbb{R}^N)\,|\,u(t_i)=P_i,\ i=0,1\}$$

使 I 达到极小, 其中 PWC^1 是逐段 C^1 的连续函数类. 若存在 $t^*\in(t_0,t_1)$, 使得 $\dot{u}_0(t^*-0)\neq\dot{u}_0(t^*+0)$. 试证:

$$L_{p_i}(t^*,u_0(t^*),\dot{u}_0^*(t-0))=L_{p_i}(t^*,u_0(t^*),\dot{u}_0^*(t+0)),\quad i=1,\cdots,N.$$

3. 设 $L\in C^2(\mathbb{R}^N\times\mathbb{R}^N)$, J 是一个闭区间. 又设 $u\in C^2(J,\mathbb{R}^N)$ 是泛函 $I(u)=\int_J L(u(t),\dot{u}(t))dt$ 的 E-L 方程的解. 定义 Hamilton 量 $H(u,p)=\sum_{i=1}^N p_iL_{p_i}(u,p)-L(u,p)$. 求证:

$$H(u(t),\dot{u}(t))=\text{const}.$$

在有 n 个自由度的质点组, 设位置坐标为 $q=(q_1,\cdots,q_n)$, 动能为

$$T=\frac{1}{2}\sum_{i,j=1}^n a_{ij}\dot{q}_i\dot{q}_j,$$

其中 (a_{ij}) 是一个正定矩阵, 位能为

$$V=V(q_1,\cdots,q_n),$$

Lagrange 作用量为 $L=T-V$. 问: 这时 Hamilton 量 H 是什么? 上述结论的物理意义是什么?

4. (第二 Erdmann 角点条件) 在题 2 的假设, 求证:

$$\begin{aligned}&L(t^*,u_0(t^*),\dot{u}_0^*(t-0))-\sum_{i=1}^N L_{p_i}(t^*,u_0(t^*),\dot{u}_0^*(t-0))\\&\quad=L(t^*,u_0(t^*),\dot{u}_0^*(t+0))-\sum_{i=1}^N L_{p_i}(t^*,u_0(t^*),\dot{u}_0^*(t+0)),\quad i=1,\cdots,N.\end{aligned}$$

提示: 引入坐标变换

$$t = v_{N+1}(s), u_i(t) = v_i(s), \quad 1 \leqslant i \leqslant N,\ s \in \Lambda,$$

其中 $v_{N+1} : \Lambda \to J$ 是同胚, 以及函数 $F \in C^1(\mathbb{R}^{N+1} \times \mathbb{R}^N)$:

$$F(y_1, \cdots, y_{N+1}, q_1, \cdots, q_{N+1}) = L\left(y_{N+1}, y_1, \cdots, y_N, \frac{q_1}{q_{N+1}}, \cdots, \frac{q_N}{q_{N+1}}\right) q_{N+1}.$$

求证:

(1) 泛函 $K(v) = \int_\Lambda F(v(s), \dot{v}(s))ds$ 与泛函 $I(u) = \int_J L(t, u(t), \dot{u}(t))dt$ 有相同的极值, 且其极值函数可以通过上述变换相互转化.

(2)

$$\forall\, \lambda > 0, \quad F(y, \lambda q) = \lambda\, F(y, q).$$

(3) 利用正齐次性, 有如下 Euler 等式:

$$F(y, q) = \sum_{i=1}^{N+1} F_{q_i}(y, q) q_i.$$

第三讲　泛函极值的必要条件与充分条件

---

§3.1　函数极值的再回顾

设 $f \in C^2(\Omega, \mathbb{R}^1), \Omega \subset \mathbb{R}^n$ 是一个开集, $x_0 \in \Omega,\ \nabla f(x_0) = 0.$

问: x_0 成为 f 的一个极小点的必要条件是什么? 充分条件是什么?

因为 $\exists x_0$ 的一个邻域 $U \subset \Omega$ 使得

$$f(x) - f(x_0) \geqslant 0, \quad \forall\, x \in U,$$

所以 $\exists\, \varepsilon_0 > 0$, 使得当 $0 < |\varepsilon| < \varepsilon_0$ 时, $\forall h \in \mathbb{R}^N \backslash \{0\}$, $x_0 + \varepsilon h \in U$, 于是

$$f(x_0 + \varepsilon h) \geqslant f(x_0), \quad |\varepsilon| < \varepsilon_0.$$

这表明一元函数 $\varepsilon \mapsto f(x_0 + \varepsilon h)$ 以 0 为极小点. 于是有

$$\frac{d^2}{d\varepsilon^2} f(x_0 + \varepsilon h)|_{\varepsilon=0} \geqslant 0.$$

这就是

$$(d^2 f(x_0)h, h) = \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} f(x_0) h_i h_j \geqslant 0,$$

即矩阵

$$d^2 f(x_0) = \frac{\partial^2}{\partial x_i \partial x_j} f(x_0)$$

是非负定的.

反过来, 如果 $d^2 f(x_0)$ 是正定的, 那么 x_0 是 f 的一个严格极小点.

§3.2 二阶变分

现在回到泛函的极小值问题. 我们知道 E-L 方程是一阶变分的条件, 它只是极小的一个必要条件, 并不充分. 从泛函分析和微分拓扑角度看, 满足 E-L 方程的解 u_0 只是泛函的临界点 (critical point). 和有限维极值点相似, 还需要看二阶变分才能判断它是否成为极小.

设 $L \in C^2(J \times \mathbb{R}^N \times \mathbb{R}^N)$.

$$I(u) = \int_J L(t, u(t), \dot{u}(t))dt.$$

又设 $u_0 \in M$ 是泛函 I 的 E-L 方程的解:

$$E_L(u_0) = 0.$$

现在 $\forall \varphi \in C_0^\infty(J, \mathbb{R}^N)$, 如果令

$$g(s) = I(u_0 + s\varphi),$$

那么一元函数 $s \mapsto g(s)$ 便以 0 为极小点.

我们把

$$\begin{aligned}
\delta^2 I(u_0, \varphi) &= \ddot{g}(0) \\
&= \frac{d^2}{ds^2} I(u_0 + s\varphi)|_{s=0} \\
&= \frac{d^2}{ds^2} \int_J L(t, u_0(t) + s\varphi(t), \dot{u}_0(t) + s\dot{\varphi}(t))dt|_{s=0} \\
&= \sum_{i,j} \int_J [L_{u^i u^j}(t, u_0(t), \dot{u}_0(t))\varphi^i(t)\varphi^j(t) \\
&\quad +2L_{u^i p^j}(t, u_0(t), \dot{u}_0(t))\varphi^i(t)\dot{\varphi}^j(t) \\
&\quad +L_{p^i p^j}(t, u_0(t), \dot{u}_0(t))\dot{\varphi}^i(t)\dot{\varphi}^j(t)]dt
\end{aligned}$$

称为 I 在 u_0 沿 φ 的二阶变分.

一方面, 如果 u_0 是极小点, 那么必然有 $\ddot{g}(0) \geqslant 0$, 即

$$\delta^2 I(u_0, \varphi) \geqslant 0, \quad \forall \varphi \in C_0^1(J, \mathbb{R}^N). \tag{3.1}$$

另一方面, 若 $u_0 \in C_0^1(J, \mathbb{R}^N)$ 满足 E-L 方程, 并且存在 $\lambda > 0$, 使得

$$\delta^2 I(u_0, \varphi) \geqslant \lambda \int_J \{|\varphi|^2 + |\dot{\varphi}|^2\}dt, \quad \forall \varphi \in C_0^1(J, \mathbb{R}^N), \tag{3.2}$$

则 u_0 是 I 的一个严格极小点.

这是因为

$$g(s)-g(0)=g(s)-g(0)-\dot{g}(0)s=\frac{s^2}{2}\ddot{g}(\theta s)=\frac{s^2}{2}[\ddot{g}(\theta s)-\ddot{g}(0)]+\frac{s^2}{2}\ddot{g}(0),$$

其中 $\theta\in(0,1)$ 依赖于 φ.

引入下列函数矩阵:

$$\begin{aligned}A&=(L_{\dot{p}_i\dot{p}_j}(t,u,p)),\\ B&=(L_{p_iu_j}(t,u,p)),\\ C&=(L_{p_ip_j}(t,u,p)),\end{aligned}$$

以及它们沿函数 $u_0(t)$ 的限制:

$$\begin{aligned}A_{u_0}&=(L_{\dot{p}_i\dot{p}_j}(t,u_0(t),\dot{u}_0(t))),\\ B_{u_0}&=(L_{p_iu_j}(t,u_0(t),\dot{u}_0(t))),\\ C_{u_0}&=(L_{p_ip_j}(t,u_0(t),\dot{u}_0(t))).\end{aligned}$$

我们有

$$\delta^2I(u_0,\varphi)=\int_J[(A_{u_0}\dot{\varphi},\dot{\varphi})+2(B_{u_0}\dot{\varphi},\varphi)+(C_{u_0}\varphi,\varphi)]dt.$$

因为

$$\ddot{g}(s)=\int_J[(A_{u_0+s\varphi}\dot{\varphi},\dot{\varphi})+2(B_{u_0+s\varphi}\dot{\varphi},\varphi)+(C_{u_0+s\varphi}\varphi,\varphi)]dt,$$

而且 $L\in C^2$, 所以对于 $\|\varphi\|_{C^1(J)}\leqslant 1$ 当 $s\to 0$, 一致地有

$$|A_{u_0+s\varphi}-A_{u_0}|+|B_{u_0+s\varphi}-B_{u_0}|+|C_{u_0+s\varphi}-C_{u_0}|=o(1),$$

即得

$$\ddot{g}(s)-\ddot{g}(0)=o(s^2)\int_J[|\nabla\varphi|^2+|\varphi|^2]dt.$$

从而 $\forall\,\varphi\in C_0^1(J,\mathbb{R}^N)$, 当 $|s|$ 充分小时, 存在 $\varepsilon<\lambda$ 使得

$$I(u_0+s\varphi)-I(u_0)\geqslant(\lambda-\varepsilon)\int_J[|\dot{\varphi}|^2+|\varphi|^2]dt.$$ □

固然 (3.1) 和 (3.2) 分别是泛函取极小值的必要条件与充分条件, 但因其中还带有任意函数 φ, 所以并不是我们最终需要的结果.

§3.3 Legendre-Hadamard 条件

首先我们注意: 三个矩阵 A_0, B_0, C_0 在判定 u_0 成为极小点中的地位是不平等的.

事实上, $\forall\tau\in \mathrm{int}(J)$, $\forall\xi\in\mathbb{R}^N$, $\forall\mu>0$ 充分小, 取 $v\in C^1(\mathbb{R}^1), v(s)=0$, 当 $|s|\geqslant 1$, 满足 $\int_{\mathbb{R}^1}\dot v(s)^2ds=1$. 令

$$\varphi(t)=\xi\mu v\left(\frac{t-\tau}{\mu}\right),$$

则

$$\dot\varphi(t)=\xi\dot v\left(\frac{t-\tau}{\mu}\right).$$

当 $\forall\mu>0$ 充分小时,

$$\int_J\dot\varphi_i\dot\varphi_j dt=\xi_i\xi_j\mu,$$

$$\int_J\dot\varphi_i\varphi_j dt=\xi_i\xi_j\mu^2\int_{\mathbb{R}^1}v(t)\dot v(t)dt,$$

$$\int_J\varphi_i\varphi_j dt=\xi_i\xi_j\mu^3\int_{\mathbb{R}^1}v(t)^2dt.$$

代入 (3.1), 令 $\mu\to 0$, 我们得到

$$\delta^2 I(u_0,\varphi)=\mu(A_{u_0}\xi,\xi)+o(\mu).$$

现在我们引入下列 Legendre-Hadamard 条件:

$$(A_{u_0}\xi,\xi)=\sum_{i,j=1}^N L_{\dot p_i\dot p_j}(\tau,u_0(\tau),\dot u_0(\tau))\xi^i\xi^j\geqslant 0,\quad \forall\tau\in J,\quad \forall\xi\in\mathbb{R}^N. \tag{3.3}$$

如果 $\exists\lambda>0$, 使得

$$\sum_{i,j=1}^N L_{\dot p_i\dot p_j}(\tau,u(\tau),\dot u(\tau))\xi^i\xi^j\geqslant \lambda|\xi|^2,\quad \forall\tau\in J,\quad \forall\xi\in\mathbb{R}^N, \tag{3.4}$$

那么我们称其为严格 Legendre-Hadamard 条件.

定理 3.1 设 $L\in C^2(J\times\mathbb{R}^N\times\mathbb{R}^N)$. 若 $u_0\in M$ 是 I 的一个极小点, 则 Legendre-Hadamard 条件 (3.3) 成立. 反之, 若 $u_0\in M$ 满足 E-L 方程, 而且存在一个 $\lambda>0$ 使得 (3.2) 成立, 则 u_0 是 I 的一个严格极小点.

我们说过条件 (3.2) 中含有任意函数 φ, 有必要把它换成一个不含 φ 的条件, 为此我们要建立它与严格 Legendre-Hadamard 条件 (3.4) 之间的联系.

事实上, 充分条件 (3.2) 右边积分中的项 $|\varphi|^2$ 是可以去掉的. 这是因为我们有如下引理.

引理 3.1 (Poincaré) 设 $\varphi \in C_0^1(J, \mathbb{R}^N)$, 则

$$\int_J |\varphi|^2 dt \leqslant \frac{(t_2 - t_1)^2}{2} \int_J |\dot{\varphi}|^2 dt.$$

证明 因为

$$\varphi(t) = \int_{t_0}^t \dot{\varphi}(s) ds,$$

由 Schwarz 不等式,

$$|\varphi(t)|^2 \leqslant \left(\int_{t_0}^t |\dot{\varphi}(s)| ds \right)^2 \leqslant (t - t_0) \int_J |\dot{\varphi}(s)|^2 ds.$$

积分后,

$$\int_J |\varphi(t)|^2 dt \leqslant \frac{(t_1 - t_0)^2}{2} \int_J |\dot{\varphi}|^2 dt.$$

现在我们如果把 (3.2) 右端积分中的 $|\varphi|^2$ 去掉, 换成

$$\delta^2 I(u, \varphi) \geqslant \lambda \int_J |\dot{\varphi}|^2 dt, \quad \forall \varphi \in C_0^1(J, \mathbb{R}^N), \tag{3.2$'$}$$

那么定理 3.1 仍成立.

例 3.1 设 $L = \sqrt{1 + p^2}$, $M = \{u \in C^1([0, b]) \,|\, u(0) = u(b) = 0\}$.
我们知道泛函 $I(u) = \int_0^b \sqrt{1 + \dot{u}^2(t)} dt$ 的 E-L 方程

$$\frac{d}{dt} \frac{\dot{u}}{\sqrt{1 + \dot{u}^2}} = 0$$

有解 $u = 0 \in M$.

因为

$$L_{uu} = L_{up} = L_{pu} = 0, \quad L_{pp} = \frac{1}{(1 + p^2)^{3/2}},$$

所以

$$\delta^2 I(0, \varphi) = \int_0^b \dot{\varphi}^2 dt.$$

应用定理 3.1, 我们知道: $u = 0$ 是一个极小点. □

一方面, 从二阶变分的表达式我们不难看出: 如果矩阵

$$\begin{pmatrix} A_{u_0} & B_{u_0} \\ B_{u_0} & C_{u_0} \end{pmatrix}$$

是正定的, 那么 E-L 方程的解 u_0 必是极小点. 但从下例我们可以看出这矩阵的正定性并不是极小的必要条件.

例 3.2 设 $I(u) = \int_J (\dot{u}(t)^2 - u(t)^2)dt$, 则在 $u = 0$,

$$\begin{pmatrix} A_0 & B_0 \\ B_0 & C_0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

不是半正定的.

但当 $|J| = t_1 - t_0$ 充分小时, 由 Poincaré 不等式, 仍然有

$$\delta^2 I(0, \varphi) = \int_J (\dot{\varphi}^2 - \varphi^2)dt \geqslant \left(1 - \frac{(t_1 - t_0)^2}{2}\right) \int_J \dot{\varphi}^2 dt.$$

于是 $u = 0$ 还是极小点. □

另一方面, 不难验证 (3.4) 不是 u_0 成为极小点的充分条件. 下面我们要讨论: 附加什么条件, 它才能成为充分条件?

§3.4 Jacobi 场

我们来引进 Jacobi 场的概念.

现在设 $L \in C^3$, 又设 u_0 是 E-L 的解. 沿 u_0 令

$$\Phi_{u_0}(t, \xi, \eta) = (A_{u_0}\eta, \eta) + 2(B_{u_0}\xi, \eta) + (C_{u_0}\xi, \xi), \quad \forall (\xi, \eta) \in \mathbb{R}^N \times \mathbb{R}^N.$$

我们称其为附属的 (accessory) Lagrange 函数.

设 u_0 是一个极小点, 我们来考察与附属 Lagrange 函数相关的变分积分

$$Q_{u_0}(\varphi) = \int_J \Phi_{u_0}(t, \varphi(t), \dot{\varphi}(t))\, dt, \quad \forall \varphi \in C_0^1(J, \mathbb{R}^N).$$

因为 $Q_{u_0}(\varphi) = \delta^2 I(u_0, \varphi) \geqslant 0$, $\forall \varphi \in C_0^1(J, \mathbb{R}^N)$, 而且 $Q_{u_0}(\theta) = 0$, 所以 θ 是 Q_{u_0} 的极小点.

现在我们把泛函 Q_{u_0} 的定义域扩充到 $\text{Lip}_0(J, \mathbb{R}^N)$(即端点为 0 的 Lipschitz 函数类) 上, 导出它的积分形式的 E-L 方程:

$$A_{u_0}\dot{\varphi}(t) + B_{u_0}^{\top}\varphi(t) - \int_{t_0}^{t} (B_{u_0}\dot{\varphi}(t) + C_{u_0}\varphi(t))dt = \text{const}.$$

倘若 L 沿 u_0 满足严格的 Lagendre-Hadamard 条件, 即 A_{u_0} 是正定的, 那么利用上面积分形式的 E-L 方程, 可见解 $\varphi \in C^2(J,\mathbb{R}^N)$, 并且 φ 还满足一个齐次二阶常微分方程组:

$$J_{u_0}(\varphi) = \frac{d}{dt}[A_{u_0}\dot{\varphi}(t) + B_{u_0}^{\top}\varphi(t)] - [B_{u_0}\dot{\varphi}(t) + C_{u_0}\varphi(t)] = 0, \quad t \in J.$$

我们称这个方程为 Jacobi *方程*, 并称算子 J_{u_0} 为沿 E-L 方程的解 u_0 的 Jacobi *算子*.

Jacobi 算子是一个二阶线性常微分算子, 它在变分问题中扮演着 Hesse 矩阵在函数极值中的角色.

我们称 Jacobi 方程的任意一个 C^2 解为沿轨道 $u_0(t)$ 的一个 Jacobi *场*. 所有 Jacobi 场构成一个 $2N$ 维线性空间.

定理 3.2 *设 φ_0 是沿 u_0 的一个 Jacobi 场, 则 $Q_{u_0}(\varphi_0) = 0$. 反之, 若 $\varphi_0 \in \mathrm{Lip}_0(J,\mathbb{R}^N)$ 满足 $Q_{u_0}(\varphi_0) = 0$, 而且 $Q_{u_0}(\varphi) \geqslant 0$, $\forall \varphi \in C_0^1(J,\mathbb{R}^N)$, 则 φ_0 是沿 u_0 的一个 Jacobi 场.*

证明 “$\Rightarrow$” 因为 Φ_{u_0} 对 (ξ,η) 是二次齐次的, 由 Euler 公式得到

$$2\Phi_{u_0}(t,\xi,\eta) = (\Phi_{u_0})_\xi(t,\xi,\eta)\xi + (\Phi_{u_0})_\eta(t,\xi,\eta)\eta.$$

如果 $\varphi_0 \in C_0^1([a,b],\mathbb{R}^N)$, $[a,b] \subset \mathrm{int}(J)$ 是沿 u_0 的 Jacobi 场, 那么

$$\begin{aligned}
&2\int_a^b (\Phi_{u_0})(t,\varphi_0(t),\dot{\varphi}_0(t))dt \\
&= \int_a^b [(\Phi_{u_0})_\xi(t,\varphi_0(t),\dot{\varphi}_0(t))\varphi_0(t) + (\Phi_{u_0})_\eta(t,\varphi_0(t),\dot{\varphi}_0(t))\dot{\varphi}_0(t)]dt \\
&= \int_a^b [(\Phi_{u_0})_\xi(t,\varphi_0(t),\dot{\varphi}_0(t)) - \frac{d}{dt}(\Phi_{u_0})_\eta(t,\varphi_0(t),\dot{\varphi}_0(t))]\varphi_0(t)dt \\
&= -\int_a^b J_{u_0}(\varphi_0)dt = 0.
\end{aligned}$$

再由 a,b 的任意性, 即得 $Q_{u_0}(\varphi_0) = 0$.

“$\Leftarrow$” 利用光滑函数逼近, 我们得到

$$Q_{u_0}(\varphi) \geqslant 0, \quad \forall \varphi \in \mathrm{Lip}_0(J,\mathbb{R}^N).$$

于是 φ_0 是 Q_{u_0} 的一个极小点. 由前面的推理知道: 它满足积分形式的 E-L 方程, 从而也满足微分形式的 E-L 方程: $J_{u_0}(\varphi_0) = 0$. □

引理 3.2 给定一个足够光滑的 Lagrange 函数 L, 设沿其 E-L 方程的解 u_0 满足严格的 Legendre-Hadamard 条件, 即 A_{u_0} 是正定的. 又若存在 $\mu > 0$ 使得

$$Q_{u_0}(\varphi) \geqslant \mu \int_J |\varphi|^2 dt,$$

则存在 $\lambda > 0$ 使得

$$Q_{u_0}(\varphi) \geqslant \lambda \int_J (|\dot{\varphi}|^2 + |\varphi|^2) dt.$$

从而 u_0 是

$$I(u) = \int_J L(t, u(t), \dot{u}(t)) dt$$

的一个严格极小点.

证明 令 $\langle \phi, \psi \rangle = \int_J \phi(t)\psi(t)dt$. 因为 A_{u_0} 是正定的, 所以存在 $\alpha > 0$ 使得

$$\langle A_{u_0}\dot{\varphi}, \dot{\varphi} \rangle \geqslant \alpha \int_J |\dot{\varphi}|^2 dt.$$

从

$$Q_{u_0}(\varphi) = \langle A_{u_0}\dot{\varphi}, \dot{\varphi} \rangle + 2\langle B_{u_0}\dot{\varphi}, \varphi \rangle + \langle C_{u_0}\varphi, \varphi \rangle,$$

推出存在正常数 C_1, C_2, 使得

$$\begin{aligned} \alpha \int_J |\dot{\varphi}|^2 dt &\leqslant Q_{u_0}(\varphi) + C_1 \left(\left(\int_J |\dot{\varphi}|^2 dt \right)^{1/2} \left(\int_J |\varphi|^2 dt \right)^{1/2} + \int_J |\varphi|^2 dt \right) \\ &\leqslant \frac{\alpha}{2} \int_J |\dot{\varphi}|^2 dt + Q_{u_0}(\varphi) + C_2 \int_J |\varphi|^2 dt. \end{aligned}$$

再应用条件

$$\int_J |\varphi|^2 dt \leqslant \mu^{-1} Q_{u_0}(\varphi),$$

得到

$$\int_J |\dot{\varphi}|^2 dt \leqslant \frac{2}{\alpha}(1 + C_2\mu^{-1}) Q_{u_0}(\varphi).$$

利用定理 3.1 以及 Poincaré 不等式, 可见 u_0 是 I 的一个严格极小点. □

§3.5 共轭点

定义 3.1 (共轭点) 设 u_0 是 $I(u) = \int_J L(t, u(t), \dot{u}(t))dt$ 的 E-L 方程的一个解. 称 $(a, u_0(a))$ 与 $(b, u_0(b))$ 是轨道 $(t, u_0(t))$ 上的一对共轭点, 如果存在一个沿 $u_0(t)$ 的非零 Jacobi 场 $\varphi \in C_0^1([a,b], \mathbb{R}^N)$(图 3.1).

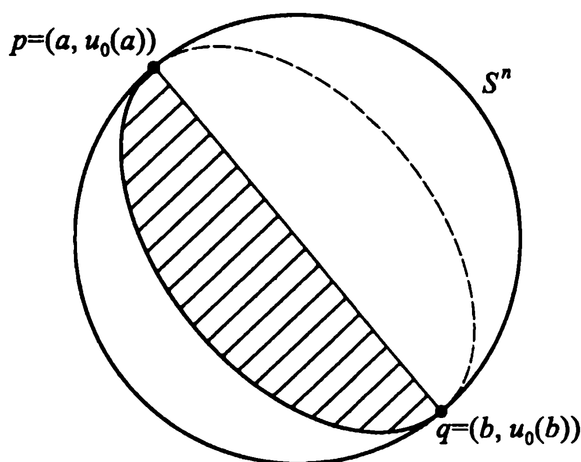


图 3.1 共轭点

有时我们把在轨道 $\{(t,u_0(t))|\,t\in(t_0,t_1]\}$ 上 $(t_0,u_0(t_0))$ 没有共轭点简称为: u_0 没有共轭点.

例 3.3 在三维欧氏空间中的二维曲面上我们有度量

$$ds^2=e(x,y)dx^2+2f(x,y)dxdy+g(x,y)dy^2.$$

我们选定其上的一条测地线 γ, 取其为 x 轴: $y=0$, 而曲线 $x=\text{const}$ 取得与测地线 γ 垂直. 引入正交曲线坐标系, 在这坐标系下, 曲线 $y=u(x)$ 上的度量表达为

$$ds^2=e(x,y)dx^2+dy^2,$$

其中 $e>0, e(x,0)=1, e_y(x,0)=0$. 弧长泛函有表达式

$$I(u)=\frac{1}{2}\int_a^b\sqrt{e(x,u)+\dot{u}^2}dx,$$

即

$$L(t,u)=\sqrt{e(x,u)+p^2}.$$

所以

$$L_{pp}=\frac{e}{(e+p^2)^{3/2}},\quad L_{up}=L_{pu}=0,\quad L_{uu}=\frac{2e_{uu}(e+p^2)-e_u^2}{4(e+p^2)^{3/2}}.$$

沿测地线 $\gamma: y=0$ 有

$$\begin{pmatrix}A_0 & B_0\\ B_0 & C_0\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & e_{uu}/2\end{pmatrix}.$$

在几何上, 把

$$K(x) = -\frac{1}{2}e_{uu}(x,0)$$

称为 Gauss 曲率. 从而附属的变分积分是

$$Q_0(\varphi) = \frac{1}{2}\int_a^b [\dot{\varphi}^2 - K(x)\varphi^2]dx.$$

Jacobi 算子是

$$J_0(\varphi) = \ddot{\varphi} + K\varphi.$$

当 K 是常数时, Jacobi 场是

$$\varphi(t) = \begin{cases} \dfrac{1}{\sqrt{-K}}\sinh\sqrt{-K}t, & K < 0, \\ t, & K = 0, \\ \dfrac{1}{\sqrt{K}}\sin\sqrt{K}t, & K > 0. \end{cases}$$

由此可见, 当 $K \leqslant 0$ 时, 无共轭点.

当 $K > 0$ 时, $(0,0)$ 的第一个共轭点是 $\left(\pi/\sqrt{K}, 0\right)$. □

注 3.1 对于一般的 Riemann 流形 (M,g), g 是 Riemann 度量. 测地线的 Lagrange 函数是

$$L(u,p) = \sum g_{ij}(u)p_i p_j,$$

对应的 Jacobi 方程是

$$\frac{d^2\varphi}{dt^2} + R\left(\dot{u}(t), \varphi(t)\right)\dot{u}(t) = 0,$$

其中 $R(\cdot,\cdot)$ 是 Riemann 曲率算子. □

定理 3.3 设 u_0 是 $I(u) = \int_J L(t,u(t),\dot{u}(t))dt$ 的 E-L 方程的一个解. 又设 A_{u_0} 是正定的. 若 $\delta^2 I(u_0,\varphi) \geqslant 0, \forall\varphi \in C_0^1(J,\mathbb{R}^N)$, 则不存在 $a \in (t_0,t_1)$ 使得 $(a,u_0(a))$ 共轭于 $(t_0,u_0(t_0))$.

证明 若不然, $\exists a \in (t_0,t_1), (a,u_0(a))$ 是 $(t_0,u_0(t_0))$ 的共轭点, 即存在沿 $u_0(t)$ 的非零 Jacobi 场 $\xi \in C^2([t_0,a],\mathbb{R}^N)$ 满足: $J_{u_0}(\xi) = 0, \xi(t_0) = \xi(a) = 0$. 令

$$\tilde{\xi}(t) = \begin{cases} \xi(t), & t \in [t_0,a], \\ 0, & t \in [a,t_1], \end{cases}$$

则 $\tilde{\xi} \in \mathrm{Lip}(J, \mathbb{R}^N)$ 满足 $\tilde{\xi}(t_0) = \tilde{\xi}(t_1) = 0$, 而且

$$Q_{u_0}(\tilde{\xi}) = \int_{t_0}^{a} \Phi_{u_0}(t, \xi(t), \dot{\xi}(t)) dt = 0.$$

根据定理 3.2, $\tilde{\xi} \in C^2(J, \mathbb{R}^N)$ 应满足 Jacobi 方程 $J_{u_0}(\tilde{\xi}) = 0$. 再由二阶微分方程初值问题的唯一性,

$$\tilde{\xi} \equiv 0.$$

这与 ξ 的非零性矛盾! □

对于 $N = 1$ 的特殊情形, 我们能证明反过来的结论.

以下设 u_0 是 E-L 方程的一个解. 我们指出: 如果在 $(t_0, t_1]$ 上 u_0 没有共轭点, 那么便存在一个正的 Jacobi 场 $\psi > 0$, $\forall t \in J$.

事实上, 设 λ 是一个 Jacobi 场, $\lambda(t_0) = 0, \dot{\lambda}(t_0) = 1$. 由假设, 它的下一个根 a 应满足: $a > t_1$. 由微分方程解对初值的连续依赖性, 存在 $\varepsilon > 0$ 和一个 Jacobi 场 ψ 满足: $\psi(t_0 - \varepsilon) = 0, \dot{\psi}(t_0 - \varepsilon) = 1, \psi(t) > 0, \forall t \in J$.

引理 3.3 设沿 u_0 有一个 Jacobi 场 $\psi(t) > 0$, $\forall t \in J$, 则 $\forall \varphi \in C_0^1(J)$ 有

$$Q_{u_0}(\varphi) = \int_J A_{u_0}(t) \psi^2(t) \left(\left(\frac{\varphi}{\psi} \right)'(t) \right)^2 dt.$$

证明 令 $\lambda = \varphi/\psi$, 则 $\varphi = \lambda\psi, \varphi' = \lambda'\psi + \lambda\psi'$. 于是

$$\begin{aligned}&A_{u_0}\varphi'^2 + 2B_{u_0}\varphi'\varphi + C_{u_0}\varphi^2 \\ &\quad = \lambda^2(A_{u_0}\psi'^2 + 2B_{u_0}\psi'\psi + C_{u_0}\psi^2) + 2\lambda'\lambda\psi(A_{u_0}\psi' + B_{u_0}\psi) + \lambda'^2 A_{u_0}\psi^2.\end{aligned}$$

由于 ψ 满足 Jacobi 方程, 所以

$$\begin{aligned}&\int_J (A_{u_0}\varphi'^2 + 2B_{u_0}\varphi'\varphi + C_{u_0}\varphi^2) dt \\ &\quad = \int_J \left\{ \left[\psi'\lambda^2(A_{u_0}\psi' + B_{u_0}\psi) + \frac{d(A_{u_0}\psi' + B_{u_0}\psi)}{dt}\psi\lambda^2 \right.\right. \\ &\qquad \left.\left. + 2\lambda'\lambda\psi(A_{u_0}\psi' + B_{u_0}\psi) \right] + A_{u_0}\lambda'^2\psi^2 \right\} dt \\ &\quad = \int_J (A_{u_0}\lambda'^2\psi^2) dt + \psi\lambda^2(A_{u_0}\psi' + B_{u_0}\psi)|_{t_0}^{t_1} \\ &\quad = \int_J (A_{u_0}\lambda'^2\psi^2) dt.\end{aligned}$$

定理 3.4 若 $N=1$, $u_0\in C^1(J)$ 是 E-L 方程的一个解. 设 $\exists\lambda>0$, 使得 $A_{u_0}(t)\geqslant\lambda$, $\forall t\in J$, 而且在 J 上存在一个正的 Jacobi 场 $\psi>0$, 则 u_0 是一个严格极小点.

证明 记

$$\alpha=\inf_J(A_{u_0}(t)\psi^2(t))>0.$$

$\forall\varphi\in C_0^1(J)$, 由引理 3.3 与 Poincaré 不等式, 立得

$$\begin{aligned}Q_{u_0}(\varphi)&=\int_J A_{u_0}\psi^2\left(\frac{\varphi}{\psi}\right)'^2dt\\&\geqslant\alpha\int_J\left(\frac{\varphi}{\psi}\right)'^2dt\\&\geqslant\alpha\frac{1}{|J|^2}\int_J\left(\frac{\varphi}{\psi}\right)^2dt\\&\geqslant\alpha\inf_J\left(\frac{1}{\psi^2}\right)\frac{1}{|J|^2}\int_J\varphi^2dt.\end{aligned}$$

所以存在 $\mu>0$ 使得

$$Q_{u_0}(\varphi)\geqslant\mu\int_J|\varphi|^2dt.$$

再应用引理 3.2 立得结论. □

例 3.4 设 $M=\{u\in C^1([0,1])\,|\,u(0)=a,u(1)=b\}$. 考虑下列泛函:

$$I(u)=\int_0^1(t\dot{u}+\dot{u}^2)dt.$$

因为 $L_u=0,L_p=2p+t,\ L_{pp}=2$, 它的 E-L 方程

$$2\ddot{u}(t)+1=0$$

有解

$$u(t)=-\frac{t^2}{4}+\left(b+\frac{1}{4}\right)t+a.$$

而其附属的变分积分是

$$Q_u(\varphi)=\int_0^1\dot{\varphi}^2dt.$$

对应的 Jacobi 方程

$$\ddot{\varphi}=0,$$

连同初始条件 $(\varphi(0),\dot{\varphi}(0))=(0,1)$ 的解是: $\varphi=t$.

这个 Jacobi 场没有共轭点, 所以 u 是一个严格极小点. □

习 题

1. 求下列泛函的极小点.

(1)

$$I(u) = \int_0^1 (t\dot{u} + \dot{u}^2)dt, \quad M = C_0^1(0,1).$$

(2)

$$I(u) = \int_a^b u\sqrt{1+\dot{u}^2}dt,\ M = \{v \in C^1[a,b] \mid v(a) = \cosh a,\ v(b) = \cosh b\},$$

其中 $0 < a < b$.

(3)

$$I(u) = \int_a^b (u^2 + \dot{u}^2)dt, \quad M = \{v \in C^1[0,b] \mid v(0) = 0,\ v(b) = B\}.$$

2. 当 φ 绝对连续, 几乎处处导数 φ' 平方可积, 并且满足 $\varphi(a) = 0$ 时, 试证 Poincaré 不等式成立:

$$\int_a^b \varphi(x)^2 dx \leqslant \frac{(b-a)^2}{2} \int_a^b \varphi'(x)^2 dx.$$

3. 在 $\mathbb{R}^3$ 中, 考虑以曲线 $r = r(z) > 0$ 绕 z 轴旋转得到的一个旋转曲面

$$S:\ r = \sqrt{x^2+y^2}.$$

(1) 求 S 上的导出度量.

(2) 写出测地线的方程.

(3) 分别就 $r = \text{const}$(圆柱), $r = z$ (圆锥) 写出测地线, 判断它们是否有极小点.

4. 设 $L = L(t,u,p)$ 可微且下方有界, 又设它关于 (u,p) 是严格凸的, $M = C_0^1(J)$, 则其 E-L 方程的解 $u \in M$ 必是泛函的严格极小点.

第四讲　强极小与极值场

§4.1　强极小与弱极小

和函数极小的概念一样, 泛函的极小也是局部的极小. 局部的概念是由邻域决定的. 变分学中的空间 M 是无穷维的函数空间, 而同一个无穷维空间经常有许多不同的拓扑, 因此有必要事先明确这个空间的拓扑是什么.

定义 4.1　设 $J=[t_0,t_1], u\in C^1(J,\mathbb{R}^N)$ 称其为

$$I=\int_J L(t,u(t),\dot{u(t)})dt$$

的强 (弱) 极小点, 如果存在 $\varepsilon>0$, 使得对于一切满足

$$\|\varphi\|_{C^0(J)}<\varepsilon \quad (\|\varphi\|_{C^1(J)}<\varepsilon)$$

的 $\varphi\in C_0^1(J,\mathbb{R}^N)$ 都有

$$I(u+\varphi)\geqslant I(u).$$

函数类 C^1 可以换成 Lip, 用 Lip 模替代 C^1 模, 在不会混淆的情况下, 也称为弱极小点. 第三讲中所说的极小点是弱极小点 (参看§3.2).

显然, 强极小点是弱极小点, Lip 意义下的弱极小点也是 C^1 意义下的弱极小点; 但反之不然. 请看下例.

例 4.1 设

$$I(u) = \int_0^1 (u'^2 + u'^3)dx,$$
$$M = \{u \in \mathrm{Lip}([0,1], \mathbb{R}^1) \mid u(0) = u(1) = 0\}, \qquad \square$$

则 $u = 0$ 是一个弱极小点.

事实上, 当 $\|u\|_{\mathrm{Lip}} < 1/2$ 时,

$$I(u) - I(0) = \int_0^1 (u'^2 + u'^3)dx \geqslant \frac{1}{2}\int_0^1 u'^2 dx \geqslant 0.$$

进一步, 从

$$\delta^2 I(0, \varphi) = \int_0^1 \dot{\varphi}^2 dt, \quad \forall \varphi \in C_0^1([0,1]),$$

与 Poincaré 不等式可知 $u = 0$ 还是一个严格 (弱) 极小点.

另一方面, 它不是强极小点. 事实上 $\forall 0 < h < 1 - h^2$, 令

$$u_h(x) = \begin{cases} -\dfrac{x}{h}, & x \in [0, h^2], \\ \dfrac{h(x-1)}{1-h^2}, & x \in [h^2, 1], \end{cases}$$

则

$$(u_h'^2 + u_h'^3)(x) = \begin{cases} \dfrac{1}{h^2} - \dfrac{1}{h^3} \leqslant -\dfrac{1}{2h^3}, & x \in [0, h^2], \\ \left(\dfrac{h}{1-h^2}\right)^2 + \left(\dfrac{h}{1-h^2}\right)^3 \leqslant 2, & x \in [h^2, 1]. \end{cases}$$

于是 $\|u_h\|_{C^0} \leqslant h$, 而且

$$I(u_h) - I(0) = I(u_h) \leqslant 2 - \frac{1}{2h} \longrightarrow -\infty \quad (h \to 0).$$

§4.2 强极小值的必要条件与 Weierstrass 过度函数

给定一个 Lagrange 函数 $L \in C^1(J \times \mathbb{R}^N \times \mathbb{R}^N)$, 设 u 是泛函

$$I(u) = \int_J L(t, u(t), \dot{u}(t))dt$$

的 E-L 方程的一个解.

我们来寻求 u 成为强极小点的必要条件. 为此我们要把 $I(u)$ 与 I 在 u 的 C^0 拓扑邻近的函数上的值作比较.

如果 u 是强极小点, 那么 $\forall\varphi\in C_0^1(J,\mathbb{R}^N)$ 满足 $\|\varphi\|_{C^0}<\varepsilon$, 有 $I(u+\varphi)\geqslant I(u)$.

现在我们来构造适当的函数 φ. $\forall\xi\in\mathbb{R}^N,\forall\tau\in(t_0,t_1)$, 当 $\lambda>0$ 充分小时, 可使得 $[\tau-\lambda^2,\tau+\lambda]\subset(t_0,t_1)$. 令

$$\psi_\lambda(s)=\begin{cases}0, & s\in(-\infty,-\lambda^2]\cup[\lambda,\infty),\\ \lambda^2+s, & s\in[-\lambda^2,0],\\ \lambda^2-\lambda s, & s\in[0,\lambda],\end{cases}$$

则

$$\psi'_\lambda(s)=\begin{cases}0, & s\in(-\infty,-\lambda^2]\cup[\lambda,\infty),\\ 1, & s\in[-\lambda^2,0],\\ -\lambda, & s\in[0,\lambda].\end{cases}$$

再令

$$\varphi_\lambda(t)=\xi\psi_\lambda(t-\tau),$$

则 $\|\varphi_\lambda\|_{C^0}=O(\lambda^2)$, $\|\dot\varphi_\lambda\|_{C^0}=\|\xi\|_{\mathbb{R}^N}$, 以及 $\|\varphi_\lambda\|_{C^1([\tau,\tau+\lambda])}=O(\lambda)$.

特别地, 取 $\varphi=\varphi_\lambda$, 我们得到

$$I(u+\varphi_\lambda)-I(u)\geqslant 0.$$

由 E-L 方程,

$$\begin{aligned}\int_J F(t)dt:=\int_J\{&L(t,u(t)+\varphi_\lambda(t),\dot u(t)+\dot\varphi_\lambda(t))-L(t,u(t),\dot u(t))\\ &-\varphi_\lambda(t)L_u(t,u(t),\dot u(t))-\dot\varphi_\lambda(t)L_p(t,u(t),\dot u(t))\}dt\\ \geqslant{}& 0,\end{aligned}$$

而且

$$\int_J F(t)dt=\left[\int_{t_0}^{\tau-\lambda^2}+\int_{\tau-\lambda^2}^{\tau}+\int_\tau^{\tau+\lambda}+\int_{\tau+\lambda}^{t_1}\right]F(s)ds,$$

其中第一与第四个积分为零, 第三个积分中的被积函数等于 $o(\lambda)$. 于是

$$\lim_{\lambda\to0}\frac{1}{\lambda^2}\int_\tau^{\tau+\lambda}F(s)ds=0.$$

而第二个积分除以 λ^2 后的极限等于

$$\lim_{\lambda\to0}\frac{1}{\lambda^2}\int_{\tau-\lambda^2}^{\tau}F(s)ds=L(\tau,u(\tau),\dot u(\tau)+\xi)-L(\tau,u(\tau),\dot u(\tau))-\xi L_p(\tau,u(\tau),\dot u(\tau)).$$

现在引入定义

$$\mathfrak{E}_L(t,u,p,q)=L(t,u,q)-L(t,u,p)-(q-p)\cdot L_p(t,u,p),$$

我们称其为 Weierstrass 过度 (excess) 函数. 它的几何意义见图 4.1.

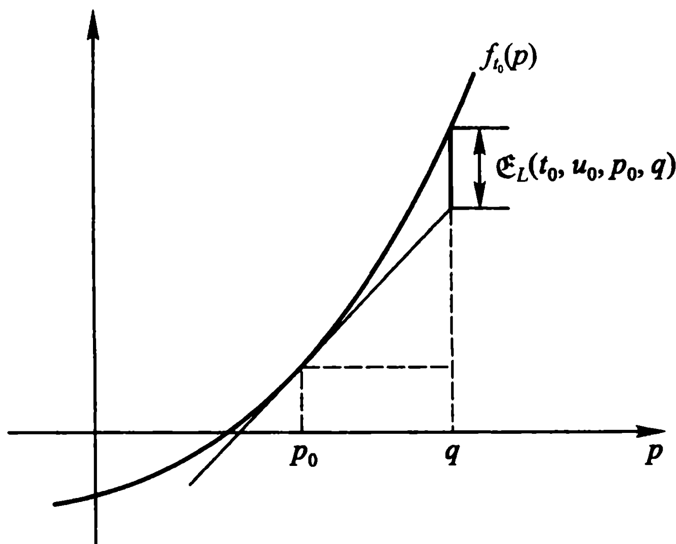


图 4.1 Weierstrass 过度函数

这里, $t_0\in J, u_0=u(t_0), p_0=\dot{u}(t_0)$, 令 $f_{t_0}(p)=L(t_0,u_0,p)$, 则 $\mathfrak{E}_L(t_0,u_0,p_0,q)$ 是曲线 f_{t_0} 在 $p=q$ 点上的值减去它在 p_0 处的切线 $f_{t_0}(p_0)+(q-p_0)\frac{d}{dp}f_{t_0}(p_0)$ 在该点上的值, 即曲线与切线的差.

总结上面的讨论, 我们得到如下定理.

定理 4.1 若 $u\in C^1(J,\mathbb{R}^N)$ 是 I 的一个强极小点, 则

$$\mathfrak{E}_L(t,u(t),\dot{u}(t),\dot{u}(t)+\xi)\geqslant 0,\quad \forall\xi\in\mathbb{R}^N,\quad \forall t\in J. \tag{4.1}$$

§4.3 极值场与强极小值

现在我们转向研究强极小值的充分条件. 为此, 对于一个给定的函数 u, 我们让它与它有共同端点的 C^0 周围的函数作比较.

习惯上, 我们把一个 E-L 方程的解 u_0 的图像 $\gamma_0=\{(t,u_0(t))\in J\times\mathbb{R}^N\,|\,t\in J\}$ 称为一条极值曲线(extremal curve).

我们要把对应于 u_0 的极值曲线嵌到它周围的极值曲线中去.

给定区间 $J=[t_0,t_1]$, Lagrange 函数 $L\in C^1(J\times\mathbb{R}^N\times\mathbb{R}^N)$, 泛函

$$I(u)=\int_J L(t,u(t),\dot{u}(t))dt$$

和 I 的 E-L 方程的一个解 u_0.

假设 u_0 可以延拓到更大的区间 $J_1 = (a, b) \supset J$ 上. 又设 $\{(t, u(t, \alpha)) \,|\, t \in J_1,\ \alpha \in B_{\varepsilon_1}(\theta) \subset \mathbb{R}^N,\ \varepsilon_1 > 0\}$ 是 I 的一族足够光滑的极值曲线.

定义 4.2 (极值场) 设 Ω 是 $\{(t, u(t, \alpha)) \,|\, t \in J,\ \alpha \in B_\varepsilon(\theta)\}$ ($\varepsilon \in (0, \varepsilon_1)$) 的一个单连通的开邻域. 又设 $\psi \in C^1(\Omega, \mathbb{R}^N)$ 是一个向量场. 如果

1.
$$\dot{u}(t) = \psi(t, u(t))$$
的每一个解 $u(t)$ 都是 I 的 E-L 方程的解;

2.
$$\det(\partial_{\alpha_i} u_j(t, \alpha)) \neq 0;$$

3.
$$\forall (t_1, u_1) \in \Omega, \exists \text{ 唯一的 } \alpha_1 \in B_{\varepsilon_1}(\theta), \text{ 使得 } u(t_1, \alpha_1) = u_1;$$

4.
$$u(t, 0) = u_0(t),$$

那么我们称 Ω 为一个极值场 (field of extremals), 并称 ψ 为其方向场 (或流) (direction field, flow) (见图 4.2).

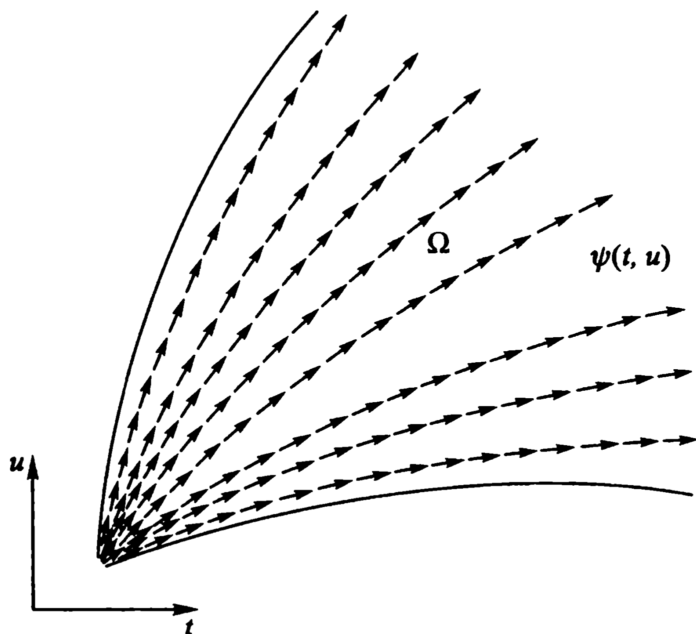


图 4.2 极值场 (Ω, ψ)

例 4.2 设 $L = \frac{1}{2}p^2$, 则其 E-L 方程 $\ddot{u} = 0$ 有解 $u_\lambda = mt + \lambda$. 因此 $\Omega = \{(t, mt + \lambda) \,|\, (t, \lambda) \in \mathbb{R}^1 \times \mathbb{R}^1\}$, $\psi(t, u) = m$ 是一个极值场和它的方向场, 其中 m 是一个常数 (图 4.3). □

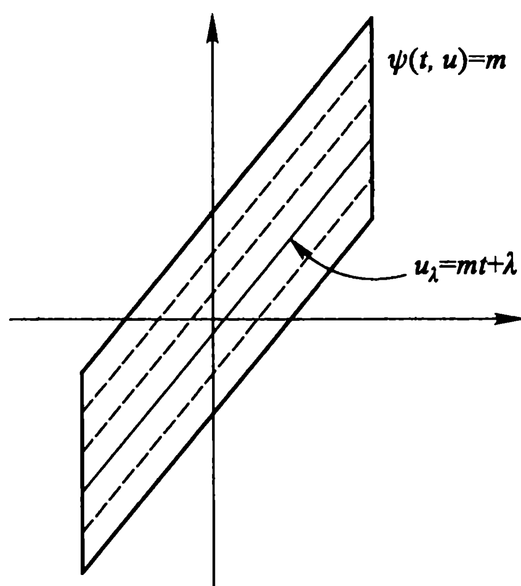


图 4.3 例 4.2

例 4.3 设 $L = \sqrt{1+p^2}$, 则其 E-L 方程 $\ddot{u} = 0$ 有解 $u_\lambda = \lambda t$. 因此 $\Omega = \{(t, \lambda t) \,|\, (t, \lambda) \in (t_0, \infty) \times \mathbb{R}^1\}$. 当 $t_0 > 0$ 时, Ω, $\psi(t, u) = u/t$ 是一个极值场和它的方向场 (见图 4.4). □

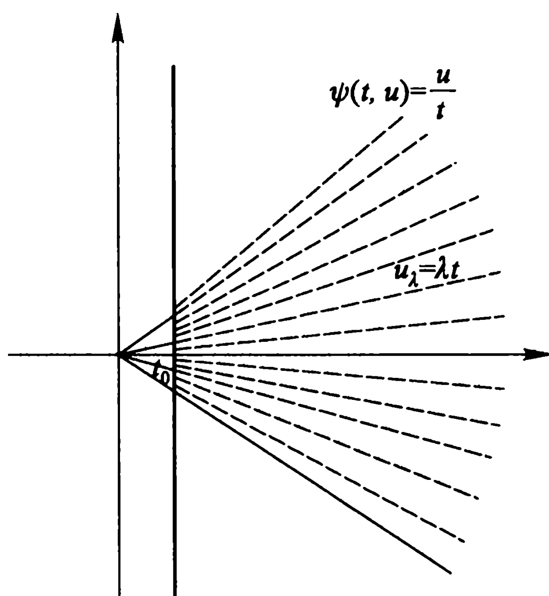


图 4.4 例 4.3

例 4.4 设 $L = \frac{1}{2}(p^2 - u^2)$, 则对应的泛函 I 的 E-L 方程 $\ddot{u} = -u$ 有解 $u_\lambda = \sin(t + \lambda), \forall \lambda \in \mathbb{R}^1$. 对于任意开区间 J, 令 $\Omega = \{(t, \sin(t + \lambda)) \,|\, (t, \lambda) \in J \times (-1, +1)\}$. 虽然极值曲线充满了整个 Ω, 但 Ω 中每一点都有两条极值曲线通过, 所以它不是极值场. □

然而, 由一族极值曲线 $u_\lambda = \lambda\sin t\,(\forall\lambda\in\mathbb{R}^1)$ 生成的 $\Omega = \{(t,\lambda\sin t)\,|\,(t,\lambda)\in(\varepsilon,\pi-\varepsilon)\times\mathbb{R}^1\}\,(\varepsilon\in(0,\pi/2))$, $\psi(t,u)=u\cot t$ 是一个极值场和它的方向场.

设极值曲线 $\gamma_0 = \{(t,u_0(t))\,|\,t\in J\}\subset\Omega$, ψ 是极值场 Ω 上的一个方向场. 我们来把 γ_0 与邻近的、端点相同的 $(u(t_0)=u_0(t_0),\ u(t_1)=u_0(t_1))$、逐段 C^1 的曲线 $\gamma=\{(t,u(t))\,|\,t\in J\}\subset\Omega$ 作比较. 注意到

$$I(u)=\int_\gamma Ldt=\int_J L(t,u(t),\dot u(t))dt,$$

而 ψ 是极值场 Ω 上的一个方向场, 所以有

$$I(u_0)=\int_{\gamma_0}[(L(t,u,\psi(t,u))-\psi(t,u)L_p(t,u,\psi(t,u)))dt+L_p(t,u,\psi(t,u))du].$$

如果积分与路径无关, 那么就有

$$\begin{aligned}I(u_0)&=\int_J L(t,u_0(t),\dot u_0(t))dt\\&=\int_\gamma[(L(t,u,\psi(t,u))-\psi(t,u)L_p(t,u,\psi(t,u)))dt+L_p(t,u,\psi(t,u))du]\\&=\int_J[L(t,u(t),\psi(t,u(t)))-(\dot u(t)-\psi(t,u(t)))L_p(t,u(t),\psi(t,u(t)))]dt.\end{aligned}$$

于是

$$\begin{aligned}I(u)-I(u_0)=\int_J&[L(t,u(t),\dot u(t))-L(t,u(t),\psi(t,u(t)))\\&-(\dot u(t)-\psi(t,u(t)))L_p(t,u(t),\psi(t,u(t)))]dt\\=\int_J&\mathfrak{E}(t,u(t),\psi(t,u(t)),\dot u(t))dt.\end{aligned}$$

如果再假定

$$\mathfrak{E}(t,u,\psi(t,u),p)\geqslant 0,\quad \forall\,(t,u,p)\in\Omega\times\mathbb{R}^N,$$

那么

$$I(u)\geqslant I(u_0),$$

即 u_0 是强极小点.

现在我们需要回答:

1. 能否把对应于 u_0 的极值曲线嵌入到一个极值场中去?

所谓 “嵌入” 是指: 存在开区间 J_1, $J\subset J_1$, 连续的 $u: J_1\times B_\varepsilon\to\mathbb{R}^N$, 使得 $\forall\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_N)\in B_\varepsilon\subset\mathbb{R}^N, u(t,\alpha)$ 是一条极值曲线, 其中 $u_0(t)=u(t,0)|_J,\forall t\in J$, 而且 $\Omega=\{(t,u(t,\alpha))\,|\,t\in J_1,\alpha\in B_\varepsilon\}$ 是一个极值场.

2. 为什么上述的积分与路径无关?

先回答第一个问题: 在什么条件下, 一条极值曲线 γ_0 能嵌入到一族极值曲线生成的极值场 Ω 中去?

为简单起见, 我们只讨论 $N=1$ 的情形.

引理 4.1 设 $L\in C^3$, $\{u(t,\alpha)\in C^2(J\times(-\varepsilon,\varepsilon))\}$ 是 E-L 方程的一族解, 则

$$\xi(t)=\frac{\partial}{\partial\alpha}u(t,\alpha)\bigg|_{\alpha=0}$$

是沿 $u_0(t)=u(t,0)$ 的一个 Jacobi 场 (图 4.5).

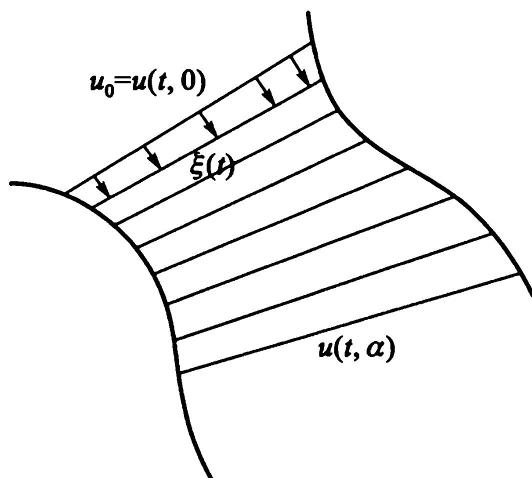


图 4.5 Jacobi 场

证明 记 $u_\alpha=u(t,\alpha)$. 对于 E-L 方程

$$\frac{d}{dt}L_p(t,u_\alpha(t),\dot{u}_\alpha(t))=L_u(t,u_\alpha(t),\dot{u}_\alpha(t)),$$

在 $\alpha=0$ 处对 α 微分, 令 $\tau=(t,u_0(t),\dot{u}_0(t))$ 得到

$$\frac{d}{dt}\Big(L_{pp}(\tau)\dot{\xi}(t)+L_{pu}(\tau)\xi(t)\Big)=L_{pu}(\tau)\dot{\xi}(t)+L_{uu}(\tau)\xi(t),$$

即

$$J_{u_0}(\xi)=0.$$ □

引理 4.2 设 $N=1, L\in C^3$, $u_0\in C^2$ 是其 E-L 方程的一个解. 又设沿 $u_0(t)$ 严格 Legendre-Hadamard 条件成立: $L_{pp}(t,u_0(t),\dot{u}_0(t))>0, \forall\, t\in(t_0,t_1]$. 如果沿对应于 u_0 的极值曲线 γ_0, 在 $(t_0,t_1]$ 上 $(t_0,u(t_0))$ 没有共轭点, 那么 γ_0 可以嵌入到一族极值曲线, 使得这族曲线确定的单连通区域 Ω 成为一个极值场.

证明 1. 利用 $L_{pp}(t,u_0(t),\dot{u}_0(t))>0,\forall\, t\in J=[t_0,t_1]$, 和 E-L 方程, 解 u_0 也可以延拓到 $J_1=(a,b)$ 上, 其中 $J\subset J_1$.

2. $\forall\alpha\in\mathbb{R}^1$, 当 $|\alpha|<\varepsilon_0$ 充分小时, 解 E-L 方程的初值问题

$$E_L(\varphi(\cdot,\alpha))=0,\ \varphi(a,\alpha)=u_0(a),\ \varphi_t(a,\alpha)=\dot{u}_0(a)+\alpha,$$

得到一族解 $\varphi(t,\alpha),\, t\in(a,b),|\alpha|<\varepsilon_0$.

由常微分方程初值问题解的唯一性, $\varphi(t,0)=u_0(t)$. 这是定义 4.2 中的 (4).

3. 现在我们定义 $\Omega_\varepsilon=\{(t,\varphi(t,\alpha))\,|\,t\in(a,b),\,|\alpha|<\varepsilon\},\varepsilon<\varepsilon_0$. 根据引理 4.1,

$$\xi(t)=\partial_\alpha\varphi(t,\alpha)|_{\alpha=0}$$

是一个沿 u_0 的 Jacobi 场, 满足

$$\xi(a)=0,\quad \dot{\xi}(a)=1.$$

按假设它没有共轭点, 利用引理 3.3 前面的说明, 可以有 $\xi(t)>0,\forall t\in[a,t_1]$. 利用常微分方程解对于参数的连续依赖性得到 $0<\varepsilon_1<\varepsilon_0$,

$$\partial_\alpha\varphi(t,\alpha)>0,\quad \forall\,|\alpha|<\varepsilon_1.$$

这是定义 4.2 中的 (2).

再应用隐函数定理, 存在 $0<\varepsilon_2<\varepsilon_1,\,\forall\,(t,u)\in\Omega_{\varepsilon_2}$, 方程

$$u=\varphi(t,\alpha)$$

存在唯一的连续可微解: $\alpha=w(t,u)\in B_{\varepsilon_2}(0)$. 这是定义 4.2 中的 (3).

4. 令 $\Omega=\Omega_{\varepsilon_2}$, 它含有 γ_0, 并且同胚于一个曲边矩形, 所以是一个单连通区域. 再看 $\varphi(t,\alpha)$ 产生的一个方向场

$$\psi(t,u)=\partial_t\varphi(t,w(t,u)),$$

则 ψ 在 Ω 内处处有定义, 并且 $\varphi(t,\alpha)$ 是方程

$$\dot{u}=\partial_t\varphi(t,\alpha)=\partial_t\varphi(t,w(t,u))=\psi(t,u)$$

的一族解曲线. 这是定义 4.2 中的 (1) (见图 4.6). □

注 4.1 当 $N>1$ 时, 引理 4.2 仍成立. 陈述如下.

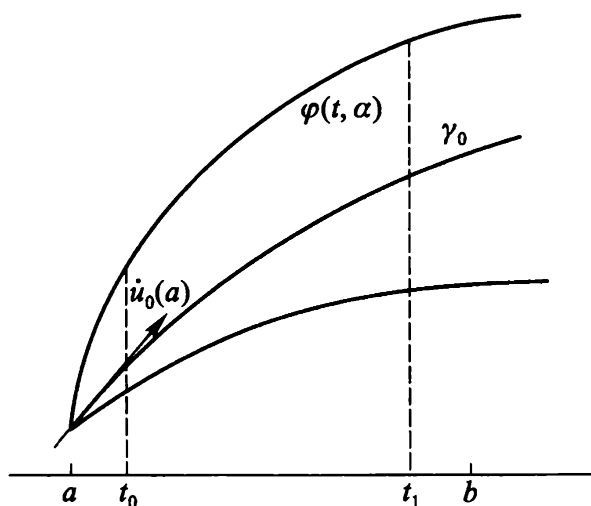


图 4.6 把 γ_0 嵌入到极值场

设 $L \in C^3([t_0,t_1]\times\mathbb{R}^N\times\mathbb{R}^N)$, $u_0 \in C^2([t_0,t_1],\mathbb{R}^N)$ 是其 E-L 方程的一个解. 又设 $\forall t\in[t_0,t_1]$, 矩阵 $L_{pp}(t,u_0(t),\dot{u}_0(t))$ 是可逆的. 如果沿对应于 u_0 的极值曲线 γ_0 在 $(t_0,t_1]$ 上 $(t_0,u(t_0))$ 没有共轭点, 那么 γ_0 可以嵌入到一族极值曲线, 使得这族曲线确定的单连通区域 Ω 成为一个极值场.

证明和引理 4.2 类似. 只需在第二段中把数 $\alpha\in\mathbb{R}^1$ 换成向量 $\alpha\in\mathbb{R}^N$. 所得 E-L 方程的解 $\varphi(t,\alpha)$ 满足

$$\partial_{\alpha_i}\varphi_j(a,\alpha)=0,\quad \partial_{\alpha_i}\partial_t\varphi_j(a,\alpha)=\delta_{ij},\quad 1\leqslant i,j\leqslant N.$$

在第三段定义 Jacobi 场为 $\{w_i(t,\alpha)=\partial_{\alpha_i}\partial_t\varphi(t,\alpha),\ i=1,\cdots,N\}$, 由于

$$w_i(a,0)=0,\quad \partial_t w_i(a,0)=e_i,\quad i=1,\cdots,N,$$

以及 γ_0 在 J_1 上没有共轭点, 得到

$$\det(\partial_{\alpha_i}\partial_t\varphi_j(t,\alpha))\neq 0,\quad \forall t\in J_1.$$

其余部分证明不变.

§4.4 Mayer 场, Hilbert 不变积分

现在我们来看第二个问题: 在什么条件下, 定理 4.1 中的积分与路径无关? 令

$$\begin{cases} R_i(t,u)=L_{p_i}(t,u,\psi(t,u)),\\ H(t,u)=\langle\psi(t,u),L_p(t,u,\psi(t,u))\rangle-L(t,u,\psi(t,u)),\end{cases}$$

以及 1-形式

$$\omega=\sum_{i=1}^N R_i du_i - H dt.$$

我们要证明:

$$d\omega = 0.$$

定义 4.3 一个极值场称为是 Mayer 场, 如果它满足下列相容性条件:

$$\partial_{u_i} L_{p_j}(t, u(t), \psi(t, u(t))) = \partial_{u_j} L_{p_i}(t, u(t), \psi(t, u(t))), \quad \forall 1 \leqslant i, j \leqslant N.$$

推论 4.1 如果 $N = 1$, 那么任意一个极值场都是 Mayer 场.

对于一个极值场和它的方向场 (Ω, ψ), 引入

$$D_\psi = \partial_t + \sum_{i=1}^N \psi_i \partial_{u_i} + \sum_{i=1}^N \left(\partial_t \psi_i + \sum_{k=1}^N \psi_k \partial_{u_k} \psi_i \right) \partial_{p_i}.$$

我们有如下引理.

引理 4.3 设 $L \in C^2(J \times \mathbb{R}^N \times \mathbb{R}^N)$. (Ω, ψ) 是一个极值场和它的方向场当且仅当对任意积分曲线 $(t, u(t))$ 都有

$$D_\psi L_p(t, u(t), \psi(t, u(t))) = L_u(t, u(t), \psi(t, u(t))).$$

证明 以下记 $\tilde{L} := \tilde{L}(t) = L(t, u(t), \psi(t, u(t)))$, 类似地记 $\tilde{L}_{u_i}, \tilde{L}_{p_i}$. (Ω, ψ) 是一个极值场和它的方向场当且仅当 $\dot{u}(t) = \psi(t, u(t))$, 而且有 E-L 方程 $\tilde{L}_u = \frac{d}{dt}\tilde{L}_p$. 直接计算:

$$\begin{aligned}
\tilde{L}_{u_i} &= \frac{d}{dt}\tilde{L}_{p_i} \\
&= \left(\partial_t + \sum_{j=1}^N \left[\psi^j \partial_{u_j} + \frac{d}{dt}\psi_j(t, u(t)) \partial_{p_j} \right] \right) \tilde{L}_{p_i} \\
&= \left(\partial_t + \sum_{j=1}^N \left[\psi^j \partial_{u_j} + \left(\partial_t \psi_j + \sum_{k=1}^N \psi_k \partial_{u_k} \psi_j \right) \partial_{p_j} \right] \right) \tilde{L}_{p_i} \\
&= D_\psi \tilde{L}_{p_i}.
\end{aligned}$$

□

引理 4.4 如果 $L \in C^2(J \times \mathbb{R}^N \times \mathbb{R}^N)$, 那么, Ω 是一个 Mayer 临界场当且仅当 $d\omega = 0$. 即

$$\partial_t R_i = -\partial_{u_i} H, \quad 1 \leqslant i \leqslant N.$$

证明 直接计算得

$$\partial_t R_i = \left(\partial_t + \sum_{j=1}^N \psi_t^j \partial_{p_j} \right) \tilde{L}_{p_i},$$

再利用相容性条件,

$$
\begin{aligned}
-H_{u_i} &= \partial_{u_i}(L-\psi L_p)(t,u,\psi(t,u)) \\
&= \left(\tilde{L}_{u_i} + \partial_{u_i}\sum_{j=1}^{N}\psi^j\partial_{p_j}\tilde{L}\right) - \partial_{u_i}\sum_{j=1}^{N}\psi^j\tilde{L}_{p_j} - \sum_{j=1}^{N}\psi^j\partial_{u_j}\tilde{L}_{p_i} \\
&= \tilde{L}_{u_i} - \sum_{j=1}^{N}\left[\psi^j\left(\partial_{u_j} + \sum_{k=1}^{N}\psi^k_{u_j}\partial_{p_k}\right)\right]\tilde{L}_{p_i}.
\end{aligned}
$$

所以

$$
\begin{aligned}
&\partial_t R_i + \partial_{u_i} H \\
&= \left(\partial_t + \sum_{i=1}^{N}\psi_t^i\partial_{p_i}\right)\tilde{L}_{p_i} + \sum_{j=1}^{N}\left(\psi^j\partial_{u_j} + \sum_{k=1}^{N}\psi^k\partial_{u_k}\psi^j\partial_{p_j}\right)\tilde{L}_{p_i} - \tilde{L}_{u_i} \\
&= D_\psi\tilde{L}_{p_i} - \tilde{L}_{u_i}.
\end{aligned}
$$

利用引理 4.3 即得 $d\omega = 0$. 反过来由 $d\omega = 0$ 以及上面的等式读者自己可以推导出相容性条件. □

给定一个 Mayer 场 (Ω, ψ), R, H 都已定义. 如果 γ 是连接 $(t_0, u(t_0))$ 与 (t, u) 的任意一条曲线, 那么线积分

$$
\begin{aligned}
S(t,u) &= \int_{(t_0,u(t_0))}^{(t,u)}\left(\sum_{i=1}^{N}R_i(t,u)du_i - H(t,u)dt\right) \\
&= \int_\gamma (L_p\cdot du + (L-\psi L_p)dt)
\end{aligned}
$$

与路径无关. 我们称这个线积分为 Hilbert 不变积分.

§4.5 强极小值的充分条件

当 $N = 1$ 时情况比较简单, 有如下定理.

定理 4.2 如果 $N = 1, L \in C^3$. 设沿其 E-L 方程的解 u_0 对应的极值曲线 γ_0 在 $(t_0, t_1]$ 没有 $(t_0, u(t_0))$ 的共轭点. 又设 $L_{pp}(t,u,p) > 0, \forall (t,u,p) \in \Omega \times \mathbb{R}^1$, 则 u_0 是 I 的一个强极小点.

证明 注意到

$$
\begin{aligned}
\mathfrak{E}(t,u,\psi(t,u),q) &= L(t,u,q) - L(t,u,\psi(t,u)) - (q-\psi(t,u))L_p(t,u,\psi(t,u)) \\
&= L_{pp}(t,u,v) \geqslant 0, \quad \forall (t,u,q) \in \Omega \times \mathbb{R}^1,
\end{aligned}
$$

其中 v 介于 $\psi(t,u)$ 与 q 之间. 应用引理 4.2 和引理 4.3 以及§4.3 中的讨论即得结论. □

例 4.5 设

$$M = \{v \in C^1([0,a]) \,|\, v(0) = (\cosh a)^{-1}, v(a) = 1\}.$$

$$I(u) = \int_0^a (\dot{u}^2 + u^2)dt$$

的 E-L 方程 $\ddot{u} - u = 0$ 有解

$$u_0 = \frac{\cosh t}{\cosh a}.$$

因为 Jacobi 方程 $\ddot{\varphi} - \varphi = 0$ 没有共轭点, 而

$$\mathfrak{E}(t,u,p,q) = (u^2 + q^2) - (u^2 + p^2) - 2p(q-p) = (q-p)^2 \geqslant 0.$$

所以 u_0 是强极小点. □

例 4.6 求泛函

$$I(u) = \int_0^1 \dot{u}^2(1+\dot{u})^2 dt$$

在边值条件 $u(0) = 0, u(1) = m$ 下的弱极小点与强极小点.

显然 I 有最小值 0. 实现它的 E-L 的解必须满足 $\dot{u} = 0$ 或 $\dot{u} = -1$.

除非 $m = 0, u(t) = 0$, 或 $m = -1, u(t) = -t$, 不可能有其他 C^1 解.

但在 Lip 函数类中, 如有解则必须 $-1 \leqslant m \leqslant 0$. 它们是

$$u_1 = \begin{cases} -t, & 0 \leqslant t \leqslant -m, \\ m, & -m \leqslant t \leqslant 1, \end{cases}$$

或

$$u_2 = \begin{cases} 0, & 0 \leqslant t \leqslant 1+m, \\ 1+m-t, & 1+m \leqslant t \leqslant 1. \end{cases}$$

计算 Lagrange 函数的二阶导数:

$$L_{uu} = L_{up} = 0, \quad L_{pp} = 2(6p^2 + 6p + 1) = 12(p-p_1)(p-p_2),$$

$L_{pp} = 0$ 有根

$$p_1 = -\frac{1}{2} - \frac{\sqrt{3}}{6}, \quad p_2 = -\frac{1}{2} + \frac{\sqrt{3}}{6}.$$

所以当 $p \notin [p_1, p_2]$ 时, $L_{pp} > 0$. 如今, 无论是前面的两个 C^1 解还是 u_1, u_2, 它们的斜率是 0 或 -1, 都在这区间之外, 因此 $L_{pp}(t, u(t), \dot{u}(t)) > 0$.

它们对应的 Jacobi 算子

$$J_u(\varphi) = 2(6p^2 + 6p + 1)\ddot{\varphi}$$

也都没有共轭点, 所以 u_1, u_2 都是弱极小点.

最后再看 Weierstrass 过度函数:

$$\begin{aligned}\mathfrak{E}(t,u,p,q) &= q^2(1+q)^2 - p^2(1+p)^2 - (q-p)(4p^3+6p^2+2p) \\ &= [q(1+q) - p(1+p)]^2 + 2p(1+p)(q-p)^2 \\ &\geqslant 0.\end{aligned}$$

它们还是强极小点. □

借助于 Mayer 场的概念, 对于高维 $N > 1$ 情形有如下定理.

定理 4.3 如果 I 的 E-L 方程的解 u_0 对应的极值曲线 γ_0 能嵌入到一族极值曲线中去, 使得这族曲线定义一个 Mayer 场 (Ω, ψ), 又设

$$\mathfrak{E}(t, u, \psi(t,u), p)) \geqslant 0, \quad \forall (t,u,p) \in \Omega \times \mathbb{R}^N,$$

则 u_0 是 I 的一个强极小点.

事实上, 当 $N > 1$ 时, 类似于定理 4.2 的结论也是对的.

定理 4.4 设 $L \in C^3(J \times \mathbb{R}^N \times \mathbb{R}^N, \mathbb{R}^1)$. 如果

(1) 沿临界曲线 $\gamma_0, (t_0, u_0(t_0))$ 没有共轭点,

(2) $(L_{p_i p_j}(t, u_0(t), \dot{u}_0(t))), \forall t \in J$ 是正定的,

(3) $\mathfrak{E}(t, u, \psi(t,u), p)) \geqslant 0, \forall (t,u,p) \in \Omega \times \mathbb{R}^N, p \neq \psi(t,u)$,

那么 u_0 是 I 的一个强极小点.

§4.6 定理 4.4 的证明 ($N > 1$ 的情形)*

我们已经知道: 如果 $N = 1$, 那么任意一个极值场 (Ω, ψ) 都是 Mayer 场. 然而, 当 $N > 1$ 时, 我们要寻求在什么条件下, 能把一条极值曲线嵌入到一个 Mayer 场中?

对于给定的 Lagrange 函数 L, 我们记 $\bar{L}(t, \alpha) = L(t, \varphi(t,\alpha), \dot{\varphi}(t,\alpha))$, 其中

$\dot{\varphi}(t,\alpha)=\varphi_t(t,\alpha)$. 类似地记 $\bar{L}_{u_i},\bar{L}_{p_i}$. 直接计算得

$$\begin{aligned}
d\omega &= d\left(\sum_{i=1}^N R_i du_i - Hdt\right)\\
&= d\left(\sum_{i=1}^N \bar{L}_{p_i} du^i + \left(\bar{L}-\sum_{i=1}^N \dot{\varphi}^i \bar{L}_{p_i}\right)dt\right)\\
&= d\left(\bar{L}dt + \sum_{i=1}^N \bar{L}_{p_i}(du^i - \dot{\varphi}^i dt)\right)\\
&= \sum_{k=1}^N \bar{L}_{u_k} d\varphi^k \wedge dt + \sum_{k=1}^N \bar{L}_{p_k} d\dot{\varphi}^k \wedge dt - \sum_{i=1}^N \bar{L}_{p_i} d\dot{\varphi}^i \wedge dt\\
&\quad + \sum_{i=1}^N\sum_{l=1}^N d\bar{L}_{p_i}\varphi^i_{\alpha_l} d\alpha^l\\
&= \sum_{k=1}^N\sum_{l=1}^N \bar{L}_{u_k}\varphi^k_{\alpha_l} d\alpha^l \wedge dt + \sum_{i=1}^N\sum_{l=1}^N \partial_t \bar{L}_{p_i}\varphi^i_{\alpha^l} dt\wedge d\alpha^l\\
&\quad + \sum_{m=1}^N\sum_{i=1}^N\sum_{l=1}^N \partial_{\alpha^m}\bar{L}_{p_i}\varphi^i_{\alpha^l} d\alpha^m \wedge d\alpha^l\\
&= \sum_{i=1}^N\sum_{l=1}^N (\bar{L}_{u_i} - \partial_t \bar{L}_{p_i})\varphi^i_{\alpha^l} d\alpha^l \wedge dt + \sum_{m=1}^N\sum_{i=1}^N\sum_{l=1}^N \partial_{\alpha^m}\bar{L}_{p_i}\varphi^i_{\alpha^l} d\alpha^m \wedge d\alpha^l.
\end{aligned}$$

引入记号

$$[\alpha^l,\alpha^m]=\sum_{i=1}^N(\partial_{\alpha^l}\bar{L}_{p_i}\varphi^i_{\alpha^m}-\partial_{\alpha^m}\bar{L}_{p_i}\varphi^i_{\alpha^l}).$$

我们称之为 Lagrange 括号. 利用 Lagrange 括号我们把上面的推导写成如下公式:

$$\begin{aligned}
d\omega &= d\left(\sum_{i=1}^N \bar{L}_{p_i} du^i + \left(\bar{L}-\sum_{i=1}^N \dot{\varphi}^i \bar{L}_{p_i}\right)dt\right)\\
&= \sum_{i=1}^N\sum_{l=1}^N E_L(\varphi)_i \varphi^i_{\alpha^l} d\alpha^l \wedge dt + \sum_{1\leqslant l<m\leqslant N} [\alpha^l,\alpha^m] d\alpha^l \wedge d\alpha^m. \qquad (4.1)
\end{aligned}$$

于是有

引理 4.5 设 $L\in C^3(J\times\mathbb{R}^N\times\mathbb{R}^N)$. 又设 (Ω,ψ) 是由一族极值曲线 $\varphi(t,\alpha)$ 决定的极值场, 则为了 (Ω,ψ) 是一个 Mayer 场, 必须且仅需

$$E_L(\varphi(\cdot,\alpha))=0,\quad \forall\alpha\in\mathbb{R}^N,\quad [\alpha^l,\alpha^m]=0,\quad 1\leqslant l,m\leqslant N.$$

证明 如果 (Ω, ψ) 是一个 Mayer 场, 那么由微分形式的不变性, (4.1) 式左边为零. 又因为极值曲线 $\varphi(\cdot, \alpha)$ 满足 E-L 方程, 所以右边第一项为零. 由此推出 Lagrange 括号为零.

反之, 若

$$E_L(\varphi(\cdot, \alpha)) = 0, \quad \forall \alpha \in \mathbb{R}^N, \quad \text{以及 } [\alpha^l, \alpha^m] = 0, \quad 1 \leqslant l, m \leqslant N,$$

则 (Ω, ψ) 是极值场. 又因为 ω 是闭形式, 所以 (Ω, ψ) 是 Mayer 场. □

对于 Lagrange 括号, 我们有

引理 4.6 设 $L \in C^3(J \times \mathbb{R}^N \times \mathbb{R}^N)$, 且 (Ω, ψ) 是一个极值场, 则

$$\frac{\partial}{\partial t}[\alpha^l, \alpha^m] = 0, \quad \forall l, m.$$

证明 利用极值场的条件直接计算:

$$\frac{\partial}{\partial t}[\alpha^l, \alpha^m] = \sum_{i=1}^N \left[\frac{\partial \bar{L}_{u_i}}{\partial \alpha^l} \varphi^i_{\alpha^m} + \frac{\partial \bar{L}_{p_i}}{\partial \alpha^l} \dot{\varphi}^i_{\alpha^m} - \frac{\partial \bar{L}_{u_i}}{\partial \alpha^m} \varphi^i_{\alpha^l} - \frac{\partial \bar{L}_{p_i}}{\partial \alpha^m} \dot{\varphi}^i_{\alpha^l} \right].$$

微分 $\frac{\partial \bar{L}_{u_i}}{\partial \alpha^l}, \frac{\partial \bar{L}_{u_i}}{\partial \alpha^m}, \frac{\partial \bar{L}_{p_i}}{\partial \alpha^l}, \frac{\partial \bar{L}_{p_i}}{\partial \alpha^m}$ 后, 利用 L 的二次可微性, 所有项都消去, 即得

$$\frac{\partial}{\partial t}[\alpha^l, \alpha^m] = 0.$$ □

接下来, 我们要在高维情形下加强注 4.1, 把 $J = [t_0, t_1]$ 上的临界曲线 γ_0 嵌入到一个 Mayer 场 (Ω, ψ) 中.

定理 4.5 设极值曲线 γ_0 在 J 上没有共轭点. 又设 $(L_{p_i p_j}(t, u(t), \dot{u}(t)))$ 是可逆的, $\forall t \in J$, 那么 γ_0 可以嵌入到一族极值曲线, 使得这族曲线定义一个 Mayer 场 (Ω, ψ).

证明 在注 4.1 的基础上, 我们得到了一个极值场 (Ω, ψ). 我们来证明这个极值场是一个 Mayer 场. 事实上, 由 $\partial_{\alpha_i} \varphi_j(a, \alpha) = 0$, $\forall i, j$, 推出 $[\alpha^l, \alpha^m](a, \alpha) = 0$, $\forall l, m \in [1, N]$. 应用引理 4.6, $\frac{d}{dt}[\alpha^l, \alpha^m](t, \alpha) = 0$, 所以 $[\alpha^l, \alpha^m](t, \alpha) = 0$. 这表明 (Ω, ψ) 是一个 Mayer 场. □

习 题

1. 用 Weierstrass 过度函数验证 $u_0 = 0$ 不是 $I(u) = \int_J (\dot{u}^2 + \dot{u}^3) dt$ 的强极小点.

2. 验证: $I(u)=\int_0^2\sqrt{u(1+\dot{u}^2)}$, $M=\{u\in C^1([0,2])\,|\,u(0)=2,\,u(2)=5\}$ 有一个双参数族的 E-L 方程解

$$u(t,\alpha,\beta)=\alpha+\frac{1}{\alpha}\left(\frac{t+\beta}{2}\right)^2.$$

当 $(\alpha,\beta)=(1,2)$ 时, $u(t,1,2)\in M$.

利用 $u(t,\alpha,\beta)$ 确定两个独立的 Jacobi 场.

3. 设 $I(u)=\int_a^b\sqrt{1+\dot{u}^2}dt$ 在 $C_0^1(0,b)$ 上定义, 写出包含 $u_0=0$ 的极值场 $\psi(t,u)$, 并验证它是一个强极小点.

4. 设 $I(u)=\int_1^2(\dot{u}+t^2\dot{u}^2)dt$ 以及 $M=\{v\in C^1[1,2]\,|\,v(1)=1,\ v(2)=2\}$. 验证

$$u_0(t)=-\frac{2}{t}+3$$

是一个强极小点.

5. 设 $u\in C^1(J\times\mathbb{R}^N,\mathbb{R}^N)$, 满足 $E_L(u(t,\alpha))=0,\forall t\in J,\forall\alpha\in\mathbb{R}^N$, $u(0,\alpha)=\theta$, $\partial_t u(0,\alpha)=\alpha$, $u_0=u(\cdot,0)$. 求证:

(1) $\partial_{\alpha_i}u(t,\alpha)|_{\alpha=0}(i=1,\cdots,N)$ 是 Jacobi 场.

(2) 如果 $(t,u_0(t))$ 在 J 没有共轭点, 那么 $\det(\partial_{\alpha_i}u_j(t,\alpha))\neq 0,\forall t\in J$.

6. 设 $I(u)=\int_a^b\left[\frac{1}{2}\dot{u}^2-V(u)\right]dt$.

(1) 当 $V(u)=\cos u$ 时, 写出沿其 E-L 方程的解 $u=0$ 的 Jacobi 算子与 Jacobi 场.

(2) 当 $V\in C^2(\mathbb{R}^1),V''(u)\leqslant 0$ 时, 求证: u_0 是 I 的严格弱极小点.

第五讲　Hamilton-Jacobi 理论

---

§5.1　程函与 Carathéodory 方程组

给定一个 Lagrange 函数 $L:\ J\times\mathbb{R}^N\times\mathbb{R}^N\to\mathbb{R}^1$ 及相应的泛函

$$I(u)=\int_J L(t,u(t),\dot{u}(t))dt.$$

从上一讲中我们知道, 对于一个极值场和它上面的一个方向场 (Ω,ψ), 有一个 1-形式

$$\omega=\langle L_p(t,u,\psi(t,u)),du\rangle-(\langle\psi(t,u),L_p(t,u,\psi(t,u))\rangle-L(t,u,\psi(t,u)))dt,$$

其中 $\langle\cdot,\cdot\rangle$ 是 $\mathbb{R}^N$ 上的内积.

为了 (Ω,ψ) 是一个 Mayer 场, 必须且仅需 ω 是一个闭形式.

于是在一个 Mayer 场上可以由 ω 定义出 Ω 上的一个单值函数 g:

$$g(t,u)-g(t_0,u_0)=\int_\gamma\omega,\tag{5.1}$$

其中 γ 是连接 (t_0,u_0) 与 (t,u) 的任意一条曲线. 我们把这个单值函数 g 称为程函 (eikonal).

注 5.1　程函有如下的物理解释: 沿一条极值曲线 $\gamma=(t,u(t))$, $\dot{u}(t)=$

$\psi(t, u(t))$, 对 ω 积分:

$$\begin{aligned} g(t_2, u(t_2)) - g(t_1, u(t_1)) &= \int_\gamma \omega \\ &= \int_J (\langle L_p(t, u(t), \psi(t, u(t))), du\rangle \\ &\quad -[\langle \psi(t, u(t)), L_p(t, u(t), \psi(t, u(t)))\rangle \\ &\quad -L(t, u(t), \psi(t, u(t)))])dt \\ &= \int_J \langle L_p(t, u(t), \dot{u}(t)), \dot{u}(t)\rangle dt \\ &\quad -[\langle \dot{u}(t), L_p(t, u(t), \dot{u}(t))\rangle - L(t, u(t), \dot{u}(t))]dt \\ &= \int_J L(t, u(t), \dot{u}(t))dt. \end{aligned}$$

这表明程函 g 在同一条极值曲线上两点的差等于 Lagrange 量沿这条曲线的积分. □

在光学上, Lagrange 量表示光在传播中瞬时走过的路程除以速度. 沿这条曲线的积分就等于光线从第一点传播到第二点所经历的时间. 因此, 从一点发出的一束光线的等时面 (物理上称为波阵面), 可以通过程函的等值面 $g(t, u) = \text{const}$ 来表示.

由定义直接推出极值场 (Ω, ψ) 上的程函 g 满足下列 Carathéodory 方程组:

$$\begin{cases} \nabla_u g(t, u) = L_p(t, u, \psi(t, u)), \\ \partial_t g(t, u) \ = L(t, u, \psi(t, u)) - \langle \psi(t, u), L_p(t, u, \psi(t, u))\rangle. \end{cases} \tag{5.2}$$

§5.2 Legendre 变换

设 $f \in C^2(\mathbb{R}^1, \mathbb{R}^1)$ 的导数 f' 有反函数 ψ. 记 $x = \psi(\xi)$. 称

$$f^*(\xi) = \xi x - f(x) = \xi\psi(\xi) - f \circ \psi(\xi)$$

为 f 的 Legendre 变换.

Legendre 变换可以推广到多元函数. 设 $f \in C^2(\mathbb{R}^N, \mathbb{R}^1)$. 又设梯度 $\xi = \nabla f(x)$ 有逆映射 ψ. 记 $x = \psi(\xi)$, 我们称

$$f^*(\xi) = \langle \xi, x\rangle - f(x) = \langle \xi, \psi(\xi)\rangle - f \circ \psi(\xi)$$

为多元函数 f 的 Legendre 变换.

Legendre 变换的几何意义如下: 记 G 为 f 的图像 $\{(x,y)\in\mathbb{R}^N\times\mathbb{R}^1\,|\,y=f(x)\}$, 它在一点 $P=(x,y)$ 处的切超平面为 $\{(\alpha,\beta)\,|\,\beta-f(x)=\langle\nabla f(x),\alpha-x\rangle\}$. 所以在这超平面上的任意一点 $Q=(\alpha,\beta)$ 满足

$$\beta-\langle\nabla f(x),\alpha\rangle=f(x)-\langle\nabla f(x),x\rangle,$$

即

$$\beta-\langle\xi,\alpha\rangle=-f^*(\xi).$$

$-f^*(\xi)$ 是这超平面在 β 轴上的截距 (见图 5.1).

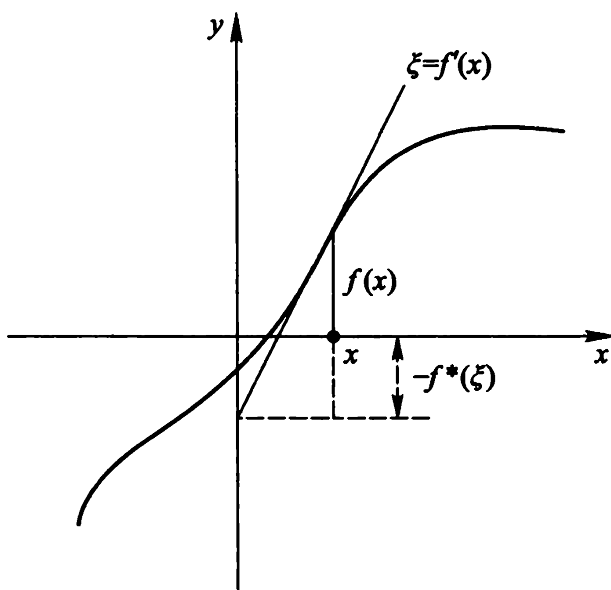


图 5.1 Legendre 变换

Legendre 变换有如下性质:

(1) 若 $f\in C^s, s\geqslant 2$, 则 $f^*\in C^s$.

证明 事实上, 因为 $\psi\in C^{s-1}$, 所以有 $f^*\in C^{s-1}$. 但由定义,

$$f^*(\xi)=\langle\xi,\psi(\xi)\rangle-f\circ\psi(\xi),$$

以及 $\xi=\nabla f(\psi(\xi))$, 可得

$$\nabla f^*(\xi)=\psi(\xi)+\langle\xi,\nabla\psi(\xi)\rangle-\langle\nabla f(\psi(\xi)),\nabla\psi(\xi)\rangle=\psi(\xi).$$

从而, $f\in C^s$. □

(2) $f^{**}=f$. 即, Legendre 变换是自反的.

事实上, 由 (1), $x=\psi(\xi)=\nabla f^*(\xi)$,

$$f^{**}(x)=\langle\xi,x\rangle-f^*(\xi)=f(x).$$

为了强调这种对称性, 我们把它写成

$$f(x)+f^*(\xi)=\langle\xi,x\rangle,\ \xi=\nabla f(x),\ x=\nabla f^*(\xi).$$

(3)

$$\left.\frac{\partial^2}{\partial\xi_i\partial\xi_j}f^*(\xi)\right|_{\xi=\nabla f(x)}=\left(\frac{\partial^2}{\partial x_i\partial x_j}\right)^{-1}f(x).$$

这是因为

$$x=\psi(\nabla f(x))=(\nabla f^*)((\nabla f)(x)),$$

等式两边再微分, 可见单位矩阵

$$I=\left.\frac{\partial^2}{\partial\xi_i\partial\xi_j}f^*(\xi)\right|_{\xi=\nabla f(x)}\cdot\frac{\partial^2}{\partial x_i\partial x_j}f(x).$$

例 5.1 设 $f(x)=x^p/p$, $p>1$, 则 $\xi=f'(x)=x^{p-1}$, 从而 $x=\xi^{\frac{1}{p-1}}$. 由定义,

$$f^*(\xi)=\xi x-f(x)=\frac{1}{p'}\xi^{\frac{1}{p'}},$$

其中 $\frac{1}{p'}+\frac{1}{p}=1$. □

例 5.2 设 $A=(a_{ij})$ 是一个对称可逆 $N\times N$ 矩阵, 又设 $f(x)=\frac{1}{2}(Ax,x)$, $x\in\mathbb{R}^N$, 则 $\xi=\nabla f(x)=Ax$ 是可逆的, 可以解出: $x=A^{-1}\xi$. 因此它的 Legendre 变换是

$$f^*(\xi)=\langle x,\xi\rangle-f(x)=\langle A^{-1}\xi,\xi\rangle-\frac{1}{2}\langle\xi,A^{-1}\xi\rangle=\frac{1}{2}\langle A^{-1}\xi,\xi\rangle.$$ □

§5.3 Hamilton 方程组

给定一个 Lagrange 函数 $L:\ \mathbb{R}^1\times\mathbb{R}^N\times\mathbb{R}^N\xrightarrow{C^2}\mathbb{R}^1$.

假设 $\det\left(L_{p_ip_j}(t,u,p)\right)\neq 0$. 我们令

$$\xi_i=L_{p_i}(t,u,p),\ 1\leqslant i\leqslant N.$$

用隐函数定理把它局部地解出来

$$p_i=\varphi_i(t,u,\xi),\ \xi=(\xi_1,\xi_2,\cdots,\xi_N),\ 1\leqslant i\leqslant N.$$

现在固定 (t,u), 作为 p 的函数, 对 L 作 Legendre 变换, 令

$$H(t,u,\xi) := L^*(t,u,\xi) = \sum_{i=1}^{N} \xi_i p_i - L(t,u,p)\big|_{p=\varphi(t,u,\xi)}. \tag{5.3}$$

我们把 H 称为 Hamilton 函数 (Hamiltonian).

因为 Legendre 变换是自反的, 既然 H 是 L 的 Legendre 变换, 那么 L 也就是 H 的 Legendre 变换.

Lagrange 函数 L 对应的 E-L 方程是二阶微分方程

$$\frac{d}{dt}L_p(t,u(t),\dot{u}(t)) - L_u(t,u(t),\dot{u}(t)) = 0,$$

也可以把它改写成等价的一阶方程组的形式:

$$\begin{cases} \dot{u}(t) = p(t), \\ \dfrac{d}{dt}L_p(t,u(t),p(t)) - L_u(t,u(t),p(t)) = 0, \end{cases}$$

其解为 $(u(t),p(t))$.

如果再设 $L \in C^3$, 那么对 (5.3) 两端作微分就有

$$\begin{aligned} & H_t dt + \langle H_u, du\rangle + \langle H_\xi, d\xi\rangle \\ &= \langle \xi, dp\rangle + \langle p, d\xi\rangle - L_t dt - \langle L_u, du\rangle - \langle L_p, dp\rangle \\ &= -L_t dt - \langle L_u, du\rangle + \langle p, d\xi\rangle. \end{aligned}$$

于是我们得到下列关系式:

$$\begin{aligned} &H_\xi(t,u,\xi) = p, && L_p(t,u,p) = \xi, \\ &H_u(t,u,\xi) + L_u(t,u,p) = 0, && H_t(t,u,\xi) + L_t(t,u,p) = 0. \end{aligned}$$

对于 E-L 方程的解 $(u(t),p(t))$. 令 $\xi(t) = L_p(t,u(t),p(t))$, 就有

$$\begin{aligned} \dot{\xi}(t) &= \frac{d}{dt}L_p(t,u(t),\dot{u}(t)) \\ &= L_u(t,u(t),\dot{u}(t)) \\ &= L_u(t,u(t),p(t)) \\ &= -H_u(t,u(t),\xi(t)) \end{aligned}$$

与

$$\dot{u}(t) = p(t) = H_\xi(t,u(t),\xi(t)).$$

即 $(u(t),\xi(t))$ 满足 Hamilton 方程组 (简称 H-S):

$$\boxed{\dot{\xi}(t) = -H_u(t,u(t),\xi(t)), \quad \dot{u}(t) = H_\xi(t,u(t),\xi(t)).} \tag{5.4}$$

反过来, 对于给定的 H-S 的解 $(u(t),\xi(t))$, 通过令 $p(t) = \dot{u}(t)$, 再由 $\xi(t) = L_p(t,u(t),p(t))$, 同样推出 $(u(t),p(t))$ 是 E-L 方程的解:

$$\frac{d}{dt}L_p(t,u(t),\dot{u}(t)) = \dot{\xi}(t) = -H_u(t,u(t),\xi(t)) = L_u(t,u(t),p(t)).$$

于是我们得到了 E-L 方程与 H-S 二者之间的一一对应关系:

$$\begin{aligned} &\text{E-L} \longleftrightarrow \text{H-S}, \\ &(u(t),p(t)) \longleftrightarrow (u(t),\xi(t)) = (u(t), L_p(t,u(t),p(t)))\,. \end{aligned}$$

Hamilton 方程组也是一个泛函的 E-L 方程. 这个泛函是

$$F(u,\xi) = \int_J [\langle \dot{u}(t),\xi(t)\rangle - H(t,u(t),\xi(t))]dt.$$

与它相应的 1-形式是

$$\alpha = \xi du - H dt,$$

我们称其为 Poincaré-Cartan 不变量.

因为 H 是 L 的 Legendre 变换, 所以 L 也是 H 的 Legendre 变换. 这就是

$$L(t,u(t),\dot{u}(t)) = \langle \dot{u}(t),\xi(t)\rangle - H(t,u(t),\xi(t)). \tag{5.5}$$

由此可见泛函 F 与泛函 I 中的被积函数实际上是同一个函数在不同变量下的表示, 而 Poincaré-Cartan 不变量实际上就是上一讲提到的 Hilbert 不变积分的微分形式.

值得注意的是: Hamilton 方程组对应的泛函 $F(u,\xi)$ 不是下方有界的, 从而不可能有最小值, H-S 的解只是相应泛函的 "临界点".

例 5.3 对于经典力学质点组 (参看第二讲例 2.1), 设 $q=(q_1,\cdots,q_N), p=(p_1,\cdots,p_N), T(p) = \frac{1}{2}\sum_{i,j=1}^N a_{ij}p_ip_j, V = V(q_1,\cdots,q_N)$. 其 Lagrange 函数是

$$L(t,q,p) = T(p) - V(q),$$

对应的 E-L 方程是

$$\frac{d}{dt}\sum_{j=1}^N a_{ij}\dot{q}_j = -\partial_{q_i}V(q), \quad i=1,\cdots,N,$$

这时 Hamilton 函数

$$H(q,\xi)=\frac{1}{2}\sum_{i,j=1}^{N}a^{ij}\xi_i\xi_j+V(q)$$

是这个质点组的能量, 其中 (a^{ij}) 是 (a_{ij}) 的逆矩阵. 对应的 Hamilton 方程组是

$$\begin{cases}\dot{\xi}_i=-\partial_{q_i}V(q),\\ \dot{q}_i=\sum\limits_{j=1}^{N}a^{ij}\xi_j, \quad i=1,\cdots,N.\end{cases}$$

当 Hamilton 函数 H 不依赖于 t 时, $\forall c\in\mathbb{R}^1$, 令 $H^{-1}(c)=\{(u,\xi)\in\mathbb{R}^N\times\mathbb{R}^N\,|\,H(u,\xi)=c\}$ 为 Hamilton 函数的一个等值面. □

定理 5.1 *Hamilton 方程组的解曲线 $\{(t,u(t),\xi(t))\,|\,\forall t\}$ 保持在同一个等值面上.*

证明 若 $(u(t),\xi(t))$ 是

$$\begin{cases}\dot{\xi}(t)=-H_u(u(t),\xi(t)),\\ \dot{u}(t)=H_\xi(u(t),\xi(t))\end{cases}$$

的解, 则由

$$\frac{d}{dt}H(u(t),\xi(t))=\langle H_u(u(t),\xi(t)),\dot{u}(t)\rangle+\langle H_\xi(u(t),\xi(t)),\dot{\xi}(t)\rangle=0,$$

立即可得 $\forall t, H(u(t),\xi(t))=\text{const}$. □

把这个定理应用到质点组, 表明: 在运动过程中能量守恒.

§5.4 Hamilton-Jacobi 方程

给定一个 Hamilton 函数 $H=H(t,u,\xi)$, 我们把一阶偏微分方程式

$$\boxed{\partial_t S(t,u)+H(t,u,\nabla_u S(t,u))=0} \tag{5.6}$$

称为 Hamilton-Jacobi 方程(简称 H-J 方程), 其中 $S=S(t,u)$ 是 $N+1$ 个变量的函数.

这个方程在经典力学和量子力学中都是一个非常基本的方程.

我们知道在一个 Mayer 场 (Ω,ψ) 上, 程函 g 满足 Carathéodory 方程组 (5.2)

$$\begin{cases}\nabla_u g(t,u)=L_p(t,u,\psi(t,u)),\\ \partial_t g(t,u)\ \ =L(t,u,\psi(t,u))-\langle\psi(t,u),L_p(t,u,\psi(t,u))\rangle.\end{cases}$$

再把 Legendre 变换 $\xi=L_p(t,u,\psi(t,u))=\nabla_u g(t,u)$ 代入 Carathéodory 方程组, 可见 g 就满足这个一阶偏微分 H-J 方程:

$$\partial_t g(t,u)+H(t,u,\nabla_u g(t,u))=0,$$

其中 H 是 Lagrange 函数 L 的 Legendre 变换.

注意到 Lagrange 函数与 Hamilton 函数之间的 Legendre 变换关系, 从给定的 Hamilton 函数 H, 我们可以写出 L, 再从 Hamilton 方程组

$$\begin{cases}\dot{\xi}_i=-H_{u_i}(t,u,\xi),\\ \dot{u}_i=H_{\xi_i}(t,u,\xi)\end{cases}\quad (1\leqslant i\leqslant n)$$

的解 $(u(t),\xi(t))$, 我们可以转换到 E-L 方程的解 $(u(t),p(t))$, 然后利用积分可以写出程函 $g(t,u)$.

由此可见, 从已知 Hamilton 方程组的一组解 $(u(t),\xi(t))$, 令

$$p(t)=H_\xi(t,u(t),\xi(t)),$$

便得到 E-L 方程的解 $(u(t),p(t))$.

再从 (5.5) 我们可以写出 Lagrange 函数 $L(t,u(t),p(t))$. 于是

$$g(t,u(t))-g(t_0,u(t_0))=\int_{t_0}^{t}L(t,u(t),p(t))dt$$

就是程函方程的解.

这表明我们可以从任取初值得到的 Hamilton 方程组的所有解导出 H-J 方程的解.

值得注意的是 Hamilton 方程组是一个常微分方程组, 而 H-J 方程是一个一阶偏微分方程.

例 5.4 (光在介质中的传播) 设在介质中一点 $(t,u)\in\mathbb{R}^1\times\mathbb{R}^N$ 的介质密度是 $\rho(t,u)$. 以真空光速为单位, 若在此点的光速为 $\frac{1}{\rho(t,u)}$, 则对应的 Lagrange 函数是

$$L(t,u,p)=\rho(t,u)\sqrt{1+p^2},$$

这时

$$\xi = L_p = \rho \cdot \frac{p}{\sqrt{1+p^2}},$$
$$p = \frac{\xi}{\sqrt{\rho^2 - \xi^2}},$$
$$H(t,u,\xi) = p\xi - L(t,u,p) = -\sqrt{\rho^2 - \xi^2}.$$

程函 g 满足程函方程

$$g_t = \sqrt{\rho^2 - |\nabla_u g|^2},$$

这是一个 H-J 方程. 有时我们把它写成

$$g_t^2 + |\nabla_u g|^2 = \rho^2$$

的形式, 而

$$g(t,u) = \text{const}$$

是光的波阵面.

对应的方向场是

$$\psi(t,u) = H_\xi(t,u,\nabla_u g) = \frac{\nabla_u g}{\sqrt{\rho^2 - |\nabla_u g|^2}} = \frac{\nabla_u g}{\partial_t g}.$$

因为

$$\dot{u} = \psi(t,u), \quad (\dot{t},\dot{u}) = (1,\dot{u}) = (\partial_t g)^{-1}(\partial_t g, \nabla_u g),$$

所以它的积分曲线在 (t,u) 空间中沿波阵面 $g(t,u) = \text{const}$ 的法方向. 也就是: “光线垂直于波阵面”. □

§5.5 Jacobi 定理*

反过来, 我们也能从 H-J 方程的解导出 H-S 的解.

定义 5.1 设 $g = g(t,u_1,\cdots,u_N;\lambda_1,\cdots,\lambda_N)$ 是 H-J 方程的一族依赖于 N 个独立参数 $(\lambda_1,\cdots,\lambda_N) \in \Lambda$ 的解 (其中 $\Lambda \subset \mathbb{R}^N$ 是一个区域). 如果 $\det(g_{u_i\lambda_j}) \neq 0$, 那么我们称它为一个完全积分.

定理 5.2 (Jacobi) 设一个 C^2 函数 $g(t,u_1,\cdots,u_N;\lambda_1,\cdots,\lambda_N)$ 是 H-J 方程的一个完全积分. 若依赖于 $2N$ 个参数 $(\alpha,\beta) = (\alpha_1,\cdots,\alpha_N,\beta_1,\cdots,\beta_N)$ 的函数

$$\begin{cases} u = U(t,\alpha,\beta), \\ p = P(t,\alpha,\beta) \end{cases}$$

满足方程

$$\begin{cases} g_{\alpha_i}(t,U(t,\alpha,\beta),\alpha) = -\beta_i, \\ P_i(t,\alpha,\beta) = g_{u_i}(t,U(t,\alpha,\beta),\alpha), \end{cases} \quad i = 1,2,\cdots,N, \tag{5.7}$$

则它们是 Hamilton 方程组的一族解.

证明 1. 先对 H-J 方程关于 α_i 微分, 得到

$$g_{t,\alpha_i} + \sum_{k=1}^{N} H_{\xi_k}(t,u,\nabla_u g) g_{\alpha_i,u_k} = 0, \quad i = 1,2,\cdots,N.$$

用 $u = U(t,\alpha,\beta)$ 代入后有

$$g_{t,\alpha_i}(t,U(t,\alpha,\beta),\alpha) + \sum_{k=1}^{N} H_{\xi_k}(t,U(t,\alpha,\beta),P(t,\alpha,\beta)) g_{\alpha_i,u_k}(t,U(t,\alpha,\beta),\alpha) = 0. \tag{5.8}$$

对 (5.7) 的第一个方程关于 t 微分, 得到

$$g_{t,\alpha_i}(t,U(t,\alpha,\beta),\alpha) + \sum_{k=1}^{N} g_{\alpha_i,u_k}(t,U(t,\alpha,\beta),\alpha) \dot{U}_k(t,\alpha,\beta) = 0. \tag{5.9}$$

(5.9) 减 (5.8) 得到

$$\sum_{k=1}^{N} [\dot{U}_k(t,\alpha,\beta) - H_{\xi_k}(t,U(t,\alpha,\beta),P(t,\alpha,\beta))] g_{\alpha_i,u_k}(t,U(t,\alpha,\beta),\alpha) = 0.$$

因为矩阵 (g_{α_i,u_k}) 是可逆的, 所以

$$\dot{U}_k(t,\alpha,\beta) = H_{\xi_k}(t,U(t,\alpha,\beta),P(t,\alpha,\beta)).$$

这是 H-S 的一组方程.

2. 再对 H-J 方程关于 u_i 微分, 得到

$$g_{t,u_i}(t,u,\alpha) + H_{u_i}(t,u,\nabla_u g(t,u,\alpha)) + \sum_{k=1}^{N} H_{\xi_k}(t,u,\nabla_u g(t,u,\alpha)) g_{u_i,u_k}(t,u,\alpha) = 0.$$

用 $u = U(t,\alpha,\beta)$, $P(t,\alpha,\beta) = \nabla_u g(t,U(t,\alpha,\beta),\alpha)$ 代入后得到

$$-H_{u_i}(t,U,P) = g_{t,u_i}(t,U,\alpha) + \sum_{k=1}^{N} g_{u_i,u_k}(t,U,\alpha)\dot{U}_k. \tag{5.10}$$

再对 (5.7) 的第二个方程关于 t 微分, 得到

$$\dot{P}_i(t,\alpha,\beta) = g_{u_i,t}(t,U,\alpha) + \sum_{k=1}^{N} g_{u_i,u_k}(t,U,\alpha)\dot{U}_k(t,\alpha,\beta), \tag{5.11}$$

即

$$\dot{P}_i(t,\alpha,\beta) = -H_{u_i}(t,U(t,\alpha,\beta),P(t,\alpha,\beta)).$$

这便导出了 H-S 的另一组方程. □

注 5.2 Jacobi 定理的意义在于: 可以用 H-J 方程的解写出 H-S 的通解, 办法是先解 N 个隐函数方程

$$g_{\alpha_i}(t,u,\alpha) = -\beta_i, \quad i = 1,\cdots,N,$$

得到

$$u = U(t,\alpha,\beta). \tag{5.12}$$

再将其代入程函 g 得到

$$p = P(t,\alpha,\beta) = \nabla_u g(t,U(t,\alpha,\beta),\alpha). \tag{5.13}$$

这一组 (u,p) 就是我们要求的 H-S 的通解. □

注 5.3 完全积分与通解的意义是不同的. 这是因为 H-J 方程是一个一阶偏微分方程, 按 Cauchy 问题解的唯一性可见它的通解中应含有一个任意函数 $\varphi(u)$, 而不仅仅是 $2N$ 个独立参数. □

但是, H-S 的通解却可以从 H-J 方程的一个完全积分 g 定出来, 也就是说, 对于给定了的任意初值 (u_0,ξ_0), H-S 的初值问题

$$\dot{u} = H_\xi(t,u,\xi), \quad \dot{\xi} = -H_u(t,u,\xi), \quad u(0) = u_0, \quad \xi(0) = \xi_0$$

的解, 能通过如下步骤求得:

给定 (u_0,ξ_0) 后, 由于 $\det(g_{u_i\alpha_j}) \neq 0$ 可以用隐函数定理对于方程组

$$\begin{cases} \nabla_\alpha g(0,u_0,\alpha) = -\beta, \\ \xi_0 = \nabla_u g(0,u_0,\alpha) \end{cases}$$

中的第二个方程先解出

$$\alpha_0 = \alpha(u_0,\xi_0).$$

再令

$$\beta_0 = -\nabla_\alpha g(0, u_0, \alpha_0),$$

代入 (5.12)-(5.13) 就得到以 (u_0, ξ_0) 为初值的 Hamilton 方程组的解.

例 5.5 (谐振子) 给定 Lagrange 函数 $L = \frac{1}{2}(mp^2 - ku^2)$, 其中 m 与 k 都为正常数. 于是 Hamilton 函数是

$$H = \frac{1}{2}\left(\frac{p^2}{m} + ku^2\right).$$

对应的 Hamilton 方程组

$$\begin{cases} \dot{u} = \dfrac{p}{m}, \\ \dot{p} = -ku \end{cases}$$

有解

$$\begin{cases} u = C\sin\sqrt{\dfrac{k}{m}}(t + t_0), \\ p = C\sqrt{mk}\cos\sqrt{\dfrac{k}{m}}(t + t_0), \end{cases}$$

其中 t_0, C 是任意常数.

我们现在利用 H-J 方程

$$g_t + H(t, u, g_u) = 0$$

与 Jacobi 定理来把这个 H-S 的解写出来.

考虑一个特殊的 $g(t, u, \alpha) = \varphi(u) - \alpha t$, 其中 α 是一个参数, 而 φ 是一个待定的函数. 于是 H-J 方程成为

$$\frac{1}{2}\left(\frac{\varphi'^2(u)}{m} + ku^2\right) = \alpha,$$

即

$$\varphi'(u) = \sqrt{m(2\alpha - ku^2)}.$$

解出来有

$$g = g(t, u, \alpha) = \int_0^u \sqrt{m(2\alpha - ku^2)}du - \alpha t.$$

现在我们来解方程

$$\beta = -g_\alpha = t - \sqrt{\frac{m}{k}}\arcsin\left(\sqrt{\frac{k}{2\alpha}}u\right),$$

得到

$$u = \sqrt{\frac{2\alpha}{k}} \sin\left(\sqrt{\frac{k}{m}}(t-\beta)\right).$$

再代入程函 g, 又得到

$$p = g_u = \varphi'(u) = \sqrt{2\alpha m} \cos\left(\sqrt{\frac{k}{m}}(t-\beta)\right).$$

这就是谐振子 Hamilton 方程组的带有两个任意参数 α, β 的解. □

习 题

1. 给定下列 Lagrange 量, 求 Hamilton 量, 并求解 Hamilton 方程组.

(1) $L = (p + ku)^2, k \neq 0,$

(2) $L = e^{-u}\sqrt{1+p^2}.$

写出 H-J 方程, 并试求其完全解.

2. 设 $\forall\,(t,u)$, $L(t,u,p)$ 对 p 是凸的, 求证:

$$H(t,u,\xi) = \sup_{p\in\mathbb{R}^N} \{\langle p,\xi\rangle - L(t,u,p)\}.$$

第六讲　含多重积分的变分问题

前面我们讨论了含单重积分的变分问题. 当未知量是多元（也还可以是向量值的）函数时, 就遇到含多重积分的变分问题. 在这一讲, 我们要把含单重积分的变分理论推广到多重积分的变分问题, 包括 E-L 方程, 弱、强极小的概念与判定, Jacobi 场以及 Weierstrass 过度函数等.

给定 $\Omega \subset \mathbb{R}^n$ 为一有界区域, 其中 $\partial\Omega \in C^1$. 给定一个 Lagrange 函数 $L = L(x, u, p) \in C^2(\bar{\Omega} \times \mathbb{R}^N \times \mathbb{R}^{nN})$, 以及在边界上的函数 $\Phi \in C^1(\partial\Omega, \mathbb{R}^N)$. 我们要在给定的边值条件 $u \in M := \{v \in C^1(\bar{\Omega}, \mathbb{R}^N)\, v|_{\partial\Omega} = \Phi|_{\partial\Omega}\}$ 下, 求下列泛函的极小值:

$$I(u) = \int_\Omega L(x, u(x), \nabla u(x))dx.$$

为了书写简便起见, 我们引入记号:

$$\begin{aligned}
&x = (x_\alpha)_1^n = (x_1, \cdots, x_n),\\
&u = (u^i)_1^N = (u^1, \cdots, u^N),\\
&p = (p_\alpha^i)_{1 \leqslant i \leqslant N, 1 \leqslant \alpha \leqslant n},\\
&p^i = (p_1^i, \cdots, p_n^i), \quad 1 \leqslant i \leqslant N,\\
&\partial_\alpha u = \frac{\partial u}{\partial x_\alpha}, \quad \alpha = 1, \cdots, n,\\
&\nabla u = \left(\frac{\partial u^i}{\partial x_\alpha}\right) = (u_\alpha^i).
\end{aligned}$$

§6.1　Euler-Lagrange 方程的推导

称 $u_0 \in M$ 是 I 在 M 上的极小点, 如果

$$I(u) \geqslant I(u_0), \quad \forall u \in U,$$

其中 $U \subset M$ 是 u_0 的一个邻域. 当取 C^1 拓扑时, u_0 称为弱极小点, 当取 C^0 拓扑时, u_0 称为强极小点.

类似于单变量的情形, 我们在假定 u_0 存在的前提下, 寻求它应满足的必要条件, 即 E-L 方程.

为简洁起见, 有时我们用记号 $\tau = (x, u_0(x), \nabla u_0(x))$.

定理 6.1　*设 $L \in C^2, u_0 \in C^2$, 并且 u_0 是 I 在 M 上的一个极小点, 则它满足下列 E-L 方程:*

$$\boxed{\sum_{\alpha=1}^{n} \frac{\partial L_{p_\alpha^i}(x, u_0(x), \nabla u_0(x))}{\partial x_\alpha} - L_{u^i}(x, u_0(x), \nabla u_0(x)) = 0, \quad 1 \leqslant i \leqslant N.} \tag{6.1}$$

证明　$\forall \varphi \in C_0^1(\bar{\Omega}, \mathbb{R}^N)$, 考虑一元函数 $s \mapsto g(s) := I(u_0 + s\varphi)$. 由 g 以 0 为极小点可推出

$$\begin{aligned} 0 = g'(0) &= \sum_{i=1}^{N} \int_\Omega \left[L_{u^i}(\tau)\varphi^i(x) + \sum_{\alpha=1}^{n} L_{p_\alpha^i}(\tau)\partial_{x_\alpha}\varphi^i(x) \right] dx \\ &= \sum_{i=1}^{N} \int_\Omega \left[L_{u^i}(\tau) - \sum_{\alpha=1}^{n} \partial_{x_\alpha} L_{p_\alpha^i}(\tau) \right] \varphi^i(x) dx. \end{aligned}$$

利用下列类似的 du Bois-Reymond 引理 (推论 6.1) 即得结论. □

我们将从更一般的情形来证明高维的 du Bois-Reymond 引理. 先引入单变量的钟形函数

$$\psi(t) = \begin{cases} \exp\left(\dfrac{-1}{1 - |t|^2}\right), & |t| < 1, \\ 0, & |t| > 1. \end{cases}$$

对于多变量 $x = (x_1, \cdots, x_n)$, 令

$$\varphi(x) = c_n^{-1} \psi(|x|),$$

其中 $|x| = (x_1^2 + \cdots + x_n^2)^{1/2}, \quad c_n = \int_{\mathbb{R}^n} \psi(|x|) dx.$

对任意充分小的 $\varepsilon > 0$, 再令

$$\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi\left(\frac{x}{\varepsilon}\right).$$

给定一个区域 $\Omega \subset \mathbb{R}^n$, 对任意充分小的 $\delta > 0$, 记 $\Omega_\delta = \{x \in \Omega \,|\, d(x, \partial\Omega) \geqslant \delta,\, |x| \leqslant 1/\delta\}$.

设 $u \in L^1_{\text{loc}}(\Omega)$, 令

$$u_\delta(x) = \int_\Omega u(y)\varphi_\delta(x-y)dy,\ \forall\, x \in \Omega_\delta.$$

作变量替换 $z = (y-x)/\delta$, 得到

$$u_\delta(x) = \int_{B_1(\theta)} u(x+\delta z)\varphi(z)dz,\ \forall\, x \in \Omega_\delta.$$

其中 $B_1(\theta)$ 是 $\mathbb{R}^n$ 中中心在原点, 半径为 1 的球. 于是 $\forall \delta_0 > 0$,

$$\int_{\Omega_{\delta_0}} |u(x)-u_\delta(x)|dx \leqslant \int_{B_1(\theta)} \varphi(z)dz \int_{\Omega_{\delta_0}} |u(x)-u(x+\delta z)|dx \to 0, \quad \text{当 } \delta \to 0.$$

事实上, 如果我们用 $\bar{\Omega}_{\delta_0}$ 上的连续函数 v 取代 u, 那么以上极限显然是成立的. 又因为 $C(\bar{\Omega}_{\delta_0})$ 在 $L^1(\Omega_{\delta_0})$ 中稠密, 所以换成 $u \in L^1_{\text{loc}}(\Omega)$ 后, 这极限还是 0.

总结如下:

引理 6.1 设 $u \in L^1_{\text{loc}}(\Omega)$, 则 $\forall\, \delta_0 > 0$,

$$\int_{\Omega_{\delta_0}} |u(x)-u_\delta(x)|dx \to 0 \quad (\delta \to 0).$$

推论 6.1 设 $u \in L^1_{\text{loc}}(\Omega)$, 又设

$$\int_\Omega u(x)\varphi(x)dx = 0, \quad \forall \varphi \in C_0^\infty(\Omega),$$

则 $u(x) = 0$ a.e..

证明 事实上, $\forall\, \delta_0 > 0$, 由于

$$\varphi_\delta(x-y) \in C_0^\infty(\Omega), \quad \forall\, x \in \Omega_{\delta_0}, \quad \forall \delta < \delta_0,$$

根据假设得到

$$u_\delta(x) = \int_\Omega u(y)\varphi_\delta(x-y)dy = 0, \quad \forall\, x \in \Omega_\delta, \quad \forall\, \delta < \delta_0.$$

再按引理 6.1, $u_\delta(x) \to u(x), \text{a.e.}\, x \in \Omega_{\delta_0}$, 这就是 $u(x) = 0, \text{a.e.}\, x \in \Omega_{\delta_0}$. 而 $\delta_0 > 0$ 是任意的, 即得 $u(x) = 0, \text{a.e.}\, x \in \Omega$. □

和一元情形一样, 我们称 $E_L : u \mapsto v = (v_1, \cdots, v_N)$, 其中

$$\boxed{v_i = \sum_{\alpha=1}^{n} \frac{\partial L_{p_\alpha^i}(x, u(x), \nabla u(x))}{\partial x_\alpha} - L_{u^i}(x, u(x), \nabla u(x)), \quad i = 1, 2, \cdots, N}$$

为关于 L 的 Euler-Lagrange 算子.

注 6.1 如果没有假设 $u \in C^2$, 那么还是可以定义 Euler-Lagrange 算子, 不过 $L_{p_\alpha^i}(x, u(x), \nabla u(x))$ 前面的 $\frac{\partial}{\partial x_\alpha}$ 应作广义微商来理解, E-L 方程也是在广义函数意义下成立的. □

类似于第二讲中注 2.2 与注 2.3, 对于含多重积分的变分问题我们也可以把泛函定义域 M 中的函数类 $C^1(\bar{\Omega}, \mathbb{R}^N)$, 换成 Lipschitz 函数类 $\mathrm{Lip}(\bar{\Omega}, \mathbb{R}^N)$, 或者更特殊些, 逐片 C^1 的连续函数类 $PWC^1(\bar{\Omega}, \mathbb{R}^N)$. $u \in PWC^1(\bar{\Omega}, \mathbb{R}^N)$ 如果存在有限个逐片 C^1 的 $n-1$ 维超曲面 $\{S_1, \cdots, S_k\}$ 使得连续的 $u \in C^1\left(\bar{\Omega} \setminus \bigcup_{j=1}^k S_j, \mathbb{R}^N\right)$, 并且 u 沿 S_j 的两侧法向导数都存在.

例 6.1 (Dirichlet 积分) 设 $N = 1, L(p) = |p|^2/2$, 对于泛函

$$D(u) = \frac{1}{2} \int_\Omega |\nabla u(x)|^2 dx,$$

有 E-L 方程

$$\Delta u = \sum_{\alpha=1}^{n} \frac{\partial^2 u}{\partial x_\alpha^2} = 0, \quad \forall x \in \Omega.$$

这就是调和方程, 又称 Laplace 方程. □

也可以考虑更一般的变分问题, 如

$$L(x, u, p) = \frac{1}{2}|p|^2 - \frac{a(x)}{\alpha+1}|u|^{\alpha+1}, \quad \alpha > 0,$$

其中 $a \in C(\bar{\Omega})$, 对应的 E-L 方程是

$$-\Delta u(x) = a(x)|u|^{\alpha-1}u, \quad \forall x \in \Omega.$$

例 6.2 (波动方程) 用 $\mathbb{R}^1 \times \mathbb{R}^3$ 表示时空连续统, 时空中一点的坐标是 (t, x), 其中 t 表示时间, $x = (x_1, x_2, x_3)$ 表示空间位置. 用 $u = u(t, x_1, x_2, x_3)$ 表示在时空区域 $\Omega \subset \mathbb{R}^1 \times \mathbb{R}^3$ 内弹性波的位移.

弹性波的动能是

$$T(u) = \frac{1}{2} \int_\Omega |\partial_t u(t, x)|^2 dtdx,$$

位能是

$$U(u)=\frac{1}{2}\int_{\Omega}|\nabla_x u(t,x)|^2dtdx,$$

Lagrange 函数是

$$I(u)=T(u)-U(u).$$

按稳定作用量原理, 弹性波的位移是这个 Lagrange 函数对应的 E-L 方程

$$\Box u=\partial_t^2u-\Delta u=0$$

的解. 这个方程就是 d'Alembert 方程. □

类似地, 如果还有内力和 (或) 外力, 那么在位能中可以再添加一些其他项. 例如:

$$U=\frac{1}{2}\int_{\Omega}(|\nabla_x u(t,x)|^2+M^2|u(t,x)|^2)dtdx,$$

其中 $M>0$ 是一个常数. 这时对应的 E-L 方程是 Klein-Gordon 方程:

$$\Box u-M^2u=\partial_t^2u-\Delta u+M^2u=0.$$

又如

$$U=\int_{\Omega}\left(\frac{1}{2}|\nabla_x u(t,x)|^2+\frac{1}{4}|u(t,x)|^4\right)dtdx,$$

这时对应的 E-L 方程是非线性波方程

$$\Box u+u^3=\partial_t^2u-\Delta u+u^3=0.$$

例 6.3 (极小曲面) 设 $\Omega\subset\mathbb{R}^n$, 给定函数 $u\in C^1(\bar{\Omega})$, 它的图像 $\{(x,u(x))|\ x\in\bar{\Omega}\}$ 是超曲面. 这个超曲面的面积是

$$A(u)=\int_{\Omega}\sqrt{1+|\nabla u(x)|^2}dx.$$

在给定边值 $u|_{\partial\Omega}=\Phi$ 的条件下, 使面积达到极小的曲面满足 E-L 方程

$$\mathrm{div}\frac{\nabla u(x)}{\sqrt{1+|\nabla u(x)|^2}}=0,\quad \forall x\in\Omega. \tag{6.2}$$

注意到这个超曲面的平均曲率是

$$H=\frac{1}{n}\mathrm{div}\frac{\nabla u(x)}{\sqrt{1+|\nabla u(x)|^2}}.$$

如果给定了平均曲率函数 $H(x)$, $x\in\Omega$, u 满足的方程是平均曲率方程

$$\operatorname{div}\frac{\nabla u(x)}{\sqrt{1+|\nabla u(x)|^2}}=nH(x),\quad \forall x\in\Omega. \tag{6.3}$$

值得注意的是, (6.3) 是泛函

$$I(u)=\int_\Omega\left(\sqrt{1+|\nabla u(x)|^2}+nH(x)u(x)\right)dx$$

的 E-L 方程. □

特别地, 当 $n=2$ 时, 方程变为

$$(1+u_y^2)u_{xx}-2u_xu_yu_{xy}+(1+u_x^2)u_{yy}=2H(x)(1+u_x^2+u_y^2)^{3/2}. \tag{6.4}$$

比较方程 (6.2) 与 (6.3), 平均曲率等于零的方程就是极小曲面的方程, 人们把平均曲率等于零的曲面称为极小曲面.

现在我们来求极小曲面方程的一个特殊解: $u(x,y)=f(x)+g(y)$. 将其代入方程 (6.3), 其中 $H=0$, 得到

$$(1+g'(y)^2)f''(x)+(1+f'(x)^2)g''(y)=0,$$

即

$$\frac{f''(x)}{1+f'(x)^2}=c=-\frac{g''(y)}{1+g'(y)^2}.$$

解出 $\arctan f'(x)=cx$, 所以

$$f(x)=-\frac{1}{c}\ln|\cos cx|.$$

同理,

$$g(y)=\frac{1}{c}\ln|\cos cy|.$$

故

$$u(x,y)=\frac{1}{c}\ln\left|\frac{\cos cy}{\cos cx}\right|.$$

这个解对应的极小曲面被称为 Scherk 曲面.

例 6.4 (Maxwell 方程) 时空中一点的坐标是 $(x^0,x^1,x^2,x^3)\in\mathbb{R}^1\times\mathbb{R}^3$, 其中 $x=(x^1,x^2,x^3)$ 是空间坐标, $x^0=ct$, t 是时间, c 是光速. □

在电磁场中电荷分布 ρ 和电流场 j 都是时空的函数. $E=(E_1,E_2,E_3), B=(B_1,B_2,B_3)$ 分别表示电场和磁场, 也是时空的函数.

Maxwell 方程组写作

$$\begin{cases} -\dfrac{1}{c}\dfrac{\partial B}{\partial t} = \nabla\times E & \text{(Faraday 定律)},\\ \nabla\cdot B = 0 & \text{(磁场的 Gauss 定律)},\\ \nabla\times B = \frac{1}{c}\frac{\partial E}{\partial t}+\frac{4\pi}{c}j & \text{(Ampere 定律)},\\ \nabla\cdot E = 4\pi\rho & \text{(电荷的 Gauss 定律)}. \end{cases}$$

因为 $\nabla\cdot B=0$, 所以有磁势 $A=(A_1,A_2,A_3)$ 使得

$$\nabla\times A = B.$$

再由 Faraday 定律,

$$\nabla\times\left(E+\frac{\partial A}{\partial x^0}\right)=\nabla\times E+\frac{1}{c}\frac{\partial B}{\partial t}=0,$$

推出存在电势 A_0 使得

$$E+\frac{\partial A}{\partial x^0}=\nabla A_0.$$

我们把 $A=(A_0,A_1,A_2,A_3)$ 称为电磁势, 并令

$$F_{ij}=\frac{\partial A_j}{\partial x^i}-\frac{\partial A_i}{\partial x^j},$$

则

$$(F_{ij})=\begin{pmatrix} 0 & -E_1 & -E_2 & -E_3\\ E_1 & 0 & B_3 & -B_2\\ E_2 & -B_3 & 0 & B_1\\ E_3 & B_2 & -B_1 & 0 \end{pmatrix}.$$

现在我们定义 Lagrange 函数

$$L=-\frac{1}{c}\left(\frac{1}{c}\sum_{i=0}^{3} j_iA_i+\frac{1}{16\pi}\sum_{i,k=0}^{3}F_{ik}^2\right),$$

其中 $J=(j_0,j_1,j_2,j_3)$, 而 $j_0=c\rho$, $j=(j_1,j_2,j_3)$.

对应的泛函是作用量

$$I(A)=\int_{\mathbb{R}^1\times\mathbb{R}^3} Ld^4x.$$

由此导出 E-L 方程

$$\frac{\partial L}{\partial A_i} - \sum_{j=0}^{3} \frac{\partial}{\partial x^j}\left(\frac{\partial L}{\partial p_i^j}\right) = 0,$$

其中

$$p_i^j = \frac{\partial A_i}{\partial x^j}.$$

因为

$$\frac{\partial L}{\partial A_i} = -\frac{1}{c^2} j_i, \ \frac{\partial L}{\partial p_i^j} = \frac{1}{4c\pi} F_{ij},$$

所以

$$\sum_{j=0}^{3} \frac{\partial F_{ij}}{\partial x^j} = -\frac{4\pi}{c} j_i, \quad i = 0, 1, 2, 3.$$

这正是 Maxwell 方程组中的 Ampere 定律 ($i = 1, 2, 3$) 与电荷的 Gauss 定律 ($i = 0$).

注 6.2 (A_0, A_1, A_2, A_3) 不是由 E, H 唯一确定的. 事实上, 对于任意给定的函数 $f \in C^1$, 令

$$A_j'(x) = A_j(x) + \frac{\partial f}{\partial x_j} \quad (j = 0, 1, 2, 3)$$

取代 $(A_j)_0^3$, 对应的 E, H 不变. 上述变换称为规范变换.

添加 Lorentz 条件

$$-\frac{\partial A_0}{\partial x_0} + \sum_{j=1}^{3} \frac{\partial A_j}{\partial x_j} = 0,$$

可以消除这种不唯一性. □

§6.2 边值条件

和单重积分的变分问题一样, 对于多重积分的变分问题, 随泛函定义域中函数在边值上的不同要求, 所得到的 E-L 方程还须添加不同的边值条件. 我们前面讨论过的定义域 M 有固定的边值函数:

$$M = \{u \in C^1(\bar{\Omega}, \mathbb{R}^N) \mid u|_{\partial\Omega} = \Phi\}.$$

如此得到的 u^* 除了满足 E-L 微分方程

$$E_L(u) = 0$$

外, 还必须满足 Dirichlet 边值条件

$$u|_{\partial\Omega} = \Phi.$$

如果把定义域换成 $M = C^1(\bar{\Omega}, \mathbb{R}^N)$, 也就是说, 在区间 Ω 的边界 $\partial\Omega$ 对泛函定义域中的函数不加任何限制, 那么类似于对单重积分的变分问题的推理, 可见泛函 I 的 C^2 极值函数 u^* 满足同样的 E-L 微分方程

$$E_L(u^*) = 0.$$

再由分部积分公式,

$$\begin{aligned}
\delta I(u^*, \varphi) &= \int_\Omega \sum_{i=1}^N \Big[L_{u^i}(x, u^*(x), \nabla u^*(x))\varphi^i(x) \\
&\quad + \sum_{\alpha=1}^n L_{p^i_\alpha}(x, u^*(x), \nabla u^*(x))\partial_\alpha \varphi^i(x) \Big] dx \\
&= \int_\Omega E_L(u^*)\dot{\varphi} + \int_{\partial\Omega} \sum_{\alpha=1}^n \sum_{i=1}^N \nu_\alpha(x) L_{p^i_\alpha}(x, u^*(x), \nabla u^*(x))\varphi^i(x) dH^{n-1}(x)
\end{aligned}$$

其中 $dH^{n-1}(x)$ 是 $\partial\Omega$ 的面积元, $\nu(x) = (\nu_1(x), \nu_2(x), \cdots, \nu_n(x))$ 是 $\partial\Omega$ 的单位外法向量, 得到如下 Neumann 边值条件:

$$\sum_{\alpha=1}^n \nu_\alpha L_{p^i_\alpha}(x, u^*(x), \nabla u^*(x))|_{\partial\Omega} = 0, \quad i = 1, 2, \cdots, N. \tag{6.5}$$

§6.3 二阶变分

给定 $\Omega \subset \mathbb{R}^n$, 给定一个 Lagrange 函数 $L \in C^2(\bar{\Omega} \times \mathbb{R}^N \times \mathbb{R}^{nN})$, 定义一个泛函 $I(u) = \int_\Omega L(x, u(x), \nabla u(x))dx$. 设 u_0 是它的 E-L 方程的一个解. 我们来研究泛函 I 的二阶变分.

对 $\forall \varphi \in C_0^\infty(\Omega, \mathbb{R}^N)$, 继续使用前面的一元函数 $g(s) = I(u_0 + s\varphi)$, 我们有

$$\begin{aligned}
\delta^2 I(u_0, \varphi) &= \ddot{g}(0) = \frac{d^2}{ds^2} I(u_0 + s\varphi)|_{s=0} \\
&= \sum_{i,j}^N \int_\Omega \Big[L_{u^i u^j}(\tau)\varphi^i(x)\varphi^j(x) + 2\sum_{\alpha=1}^n L_{u^i p^j_\alpha}(\tau)\varphi^i(x)\partial_\alpha \varphi^j(x) \\
&\quad + \sum_{\alpha,\beta=1}^n L_{p^i_\alpha p^j_\beta}(\tau)\partial_\alpha \varphi^i(x)\partial_\beta \varphi^j(x) \Big] dx,
\end{aligned}$$

其中 $\tau=(x,u_0(x),\nabla u_0(x))$.

同样引入简单记号

$$A=(L_{p_\alpha^j p_\beta^k}(x,u,p)),$$
$$B=(L_{p_\beta^j u^k}(x,u,p)),$$
$$C=(L_{u^j u^k}(x,u,p)),$$

以及

$$A_{u_0}=(a_{\alpha\beta}^{jk})=(L_{p_\alpha^j p_\beta^k}(\tau)),$$
$$B_{u_0}=(b_\beta^{jk})=(L_{p_\beta^j u^k}(\tau)),$$
$$C_{u_0}=(c^{jk})=(L_{u^j u^k}(\tau)),$$

并记

$$Q_{u_0}(\varphi)=\delta^2 I(u,\varphi)=\int_\Omega [A_{u_0}(\nabla\varphi,\nabla\varphi)+2B_{u_0}(\nabla\varphi,\varphi)+C_{u_0}(\varphi,\varphi)]dx,$$

如果 u_0 是一个弱极小点, 那么必然有

$$Q_{u_0}(\varphi)\geqslant 0,\quad \forall\varphi\in C_0^1(\Omega,\mathbb{R}^N), \tag{6.6}$$

其中 $C_0^1(\Omega,\mathbb{R}^N)$ 是 $C_0^\infty(\Omega,\mathbb{R}^N)$ 在 $C^1(\bar\Omega,\mathbb{R}^N)$ 中的闭包.

反过来, 若 $u_0\in C^1(\bar\Omega,\mathbb{R}^N)$ 满足 E-L 方程, 并且存在 $\lambda>0$ 使得

$$Q_{u_0}(\varphi)\geqslant\lambda\int_\Omega[|\nabla\varphi|^2+|\varphi|^2]dx,\quad \forall\varphi\in C_0^1(\Omega,\mathbb{R}^N), \tag{6.7}$$

则 u_0 是 I 的一个严格弱极小点. 证明和 $n=1$ 的情形一样 (见定理 3.1).

仿照 $n=1$ 的情形, 我们也有类似的 Legendre-Hadamard 条件.

$\forall x_0\in\Omega$, $\forall\mu>0$, 取 $v\in C_0^\infty(B_1(\theta),\mathbb{R}^N)$. 当 μ 充分小时, 令

$$\varphi(x)=\mu v\left(\frac{x-x_0}{\mu}\right),$$

代入 (6.6) 得到

$$\begin{aligned}Q_{u_0}(\varphi)=\mu^n\int_{B_1(\theta)}&[A_{u_0}(x_0+\mu y)\nabla v(y)\nabla v(y)\\&+2\mu B_{u_0}(x_0+\mu y)v\nabla v(y)+\mu^2 C_{u_0}(x_0+\mu y)v(y)v(y)]dy\geqslant 0.\end{aligned}$$

令 $\mu\to 0$, 得到

$$\sum_{j,k=1}^N\sum_{\alpha,\beta=1}^n a_{\alpha\beta}^{jk}(x_0)\int_{B_1(\theta)}\partial_\alpha v^j(y)\partial_\beta v^k(y)dy\geqslant 0.$$

现在对 $\forall\rho\in C_0^\infty(B_1(\theta),\mathbb{R}^1)$, 满足 $\int_{B_1(0)}\rho^2(y)dy=1$, $\forall\xi\in\mathbb{R}^N$, $\forall\eta\in\mathbb{R}^n$, 分别令

$$v(y)=\xi\cos(t\eta\cdot y)\rho(y)$$

以及

$$v(y)=\xi\sin(t\eta\cdot y)\rho(y),$$

代入以上不等式, 然后相加, 再令 $t\to\infty$ 得到

$$\sum_{j,k=1}^{N}\sum_{\alpha,\beta=1}^{n}a_{\alpha\beta}^{jk}(x_0)\xi^j\xi^k\eta_\alpha\eta_\beta+O(t^{-1})\geqslant 0,$$

即

$$\sum_{j,k=1}^{N}\sum_{\alpha,\beta=1}^{n}a_{\alpha\beta}^{jk}(x_0)\xi^j\xi^k\eta_\alpha\eta_\beta\geqslant 0.$$

这就是 Legendre-Hadamard 条件

$$\boxed{\sum_{j,k=1}^{N}\sum_{\alpha,\beta=1}^{n}L_{p_\alpha^j p_\beta^k}(x,u_0(x),\nabla u_0(x))\xi^j\xi^k\eta_\alpha\eta_\beta\geqslant 0,\ \forall(x,\xi,\eta)\in\Omega\times\mathbb{R}^N\times\mathbb{R}^n.} \tag{6.8}$$

如果我们使用秩 1 矩阵的记号

$$\pi=(\pi_\alpha^i)=(\xi^i\eta_\alpha),$$

那么 (6.8) 可以等价地写成

$$\sum_{j,k=1}^{N}\sum_{\alpha,\beta=1}^{n}L_{p_\alpha^j p_\beta^k}(x,u_0(x),\nabla u_0(x))\pi_\alpha^j\pi_\beta^k\geqslant 0,\quad\forall\pi,\quad\mathrm{rank}(\pi)=1.$$

如果 $\exists\lambda>0$ 使得

$$\sum_{j,k=1}^{N}\sum_{\alpha,\beta=1}^{n}L_{p_\alpha^j p_\beta^k}(x,u_0(x),\nabla u_0(x))\xi^j\xi^k\eta_\alpha\eta_\beta\geqslant\lambda|\xi|^2|\eta|^2,$$
$$\forall x\in\Omega,\quad\forall\xi\in\mathbb{R}^N,\quad\forall\eta\in\mathbb{R}^n, \tag{6.9}$$

那么我们称其为严格 Legendre-Hadamard 条件.

(6.9) 也可以等价地写成

$$\sum_{j,k=1}^{N}\sum_{\alpha,\beta=1}^{n}L_{p_\alpha^j p_\beta^k}(x,u_0(x),\nabla u_0(x))\pi_\alpha^j\pi_\beta^k\geqslant\lambda\|\pi\|^2,\quad\forall\pi,\quad\mathrm{rank}(\pi)=1.$$

这里矩阵 $\pi = (\pi_\alpha^j)$ 的模

$$\|\pi\| = \left(\sum_{\alpha=1}^{n}\sum_{j=1}^{N}(\pi_\alpha^j)^2\right)^{1/2}$$

对于多重积分变分问题还有一个更强的条件:

$$\sum_{j,k=1}^{N}\sum_{\alpha,\beta=1}^{n} L_{p_\alpha^j p_\beta^k}(x, u_0(x), \nabla u_0(x))\pi_\alpha^j \pi_\beta^k \geqslant \lambda\|\pi\|^2, \quad \forall x \in \Omega, \quad \forall \pi \in \mathbb{R}^{n\times N},$$

我们称其为一致强椭圆条件.

当 $N = 1$ 或 $n = 1$ 时, 严格 Legendre-Hadamard 条件与一致强椭圆条件没有区别.

总结如下.

定理 6.2 设 $L \in C^2$, 又设 $u_0 \in M$ 是 I 的一个弱极小点, 则 (6.8) 成立. 反之, 设 $u_0 \in M$ 满足 E-L 方程, 而且存在一个 $\lambda > 0$ 使得

$$\delta^2 I(u_0, \varphi) \geqslant \lambda \int_\Omega \{|\varphi(x)|^2 + |\nabla\varphi(x)|^2\}dx, \quad \forall \varphi \in C_0^1(\Omega, \mathbb{R}^N), \tag{6.10}$$

则 u_0 是 I 的一个严格弱极小点.

同理, (6.10) 蕴含了严格 Legendre-Hadamard 条件 (6.9). 但 (6.9) 并不是 u_0 成为弱极小点的充分条件.

§6.4 Jacobi 场

对于多重积分变分问题, 也有 Jacobi 场的概念.

设 $L \in C^3$, 又设 u_0 是弱极小点, 则

$$\delta^2 I(u_0, \varphi) \geqslant 0, \quad \forall \varphi \in C_0^1(\Omega, \mathbb{R}^N),$$

即 $Q_{u_0}(\varphi) \geqslant 0$.

泛函 Q_{u_0} 的 E-L 方程是一个齐次二阶偏微分方程组:

$$\sum_{k=1}^{N}\left[\sum_{\alpha=1}^{n}\partial_\alpha\left(\sum_{\beta=1}^{n} a_{\alpha\beta}^{jk}\partial_\beta\varphi^k + b_\alpha^{jk}\varphi^k\right) - \left(\sum_{\beta=1}^{n} b_\beta^{jk}\partial_\beta\varphi^k + c^{jk}\varphi^k\right)\right] = 0,$$
$$j = 1, 2, \cdots, N.$$

和 $n=1$ 的情形一样, 我们称这个方程为 Jacobi 方程, 并称

$$J_{u_0}:\varphi\mapsto(\psi^1,\cdots,\psi^N)$$

为沿 u_0 的 Jacobi 算子, 其中 $\psi^j=\sum_{k=1}^N\left[\sum_{\alpha=1}^n\partial_\alpha\left(\sum_{\beta=1}^n a_{\alpha\beta}^{jk}\partial_\beta\varphi^k+b_\alpha^{jk}\varphi^k\right)-\left(\sum_{\beta=1}^n b_\beta^{jk}\partial_\beta\varphi^k+c^{jk}\varphi^k\right)\right]$, $j=1,\cdots,N$.

Jacobi 方程的任一 C^2 解称为沿 u_0 的 Jacobi 场.

对于满足严格 Legendre-Hadamard 条件的微分算子有如下不等式:

引理 6.2 (Gårding 不等式) 设 $(a_{\alpha\beta}^{jk}(x))$ 是 $\bar\Omega\subset\mathbb{R}^n$ 上的一致连续函数, 又设存在 $\sigma>0$ 使得

$$\sum_{\alpha,\beta=1}^n\sum_{j,k=1}^N a_{\alpha\beta}^{jk}(x)\xi^j\xi^k\eta_\alpha\eta_\beta\geqslant\sigma|\xi|^2|\eta|^2,\quad\forall x\in\Omega,$$

则存在 $\alpha>0$ 及 $C_0>0$ 使得

$$\int_\Omega\sum_{\alpha,\beta=1}^n\sum_{j,k=1}^N a_{\alpha\beta}^{jk}(x)\partial_\alpha\varphi^j\partial_\beta\varphi^k dx\geqslant\alpha\int_\Omega|\nabla\varphi(x)|^2dx-C_0\int_\Omega|\varphi(x)|^2dx,$$

$\forall\varphi\in C_0^1(\Omega,\mathbb{R}^N)$.

证明 对于 $N=1$, 这个结论是显然的, 而且 $C_0=0$. 这只要从假设

$$\sum_{\alpha,\beta=1}^n a_{\alpha\beta}\partial_\alpha\varphi(x)\partial_\beta\varphi(x)\geqslant\sigma|\nabla\varphi(x)|^2$$

出发, 在不等式两边积分就得到了.

当 $N>1$, 而且 $(a_{\alpha\beta}^{jk}(x))$ 是常数时, 可以利用 Fourier 变换来证明. 让 φ 在 Ω 外定义为零, 使其在全 $\mathbb{R}^n$ 上有定义. 令

$$\hat\varphi(\xi)=\int_{\mathbb{R}^n}\varphi(x)\exp\left(-2\pi i\langle\xi,x\rangle_{\mathbb{R}^n}\right)dx,$$

则有

$$2\pi i\xi_\alpha\hat\varphi(\xi)=\int_{\mathbb{R}^n}\partial_\alpha\varphi(x)\exp\left(-2\pi i\langle\xi,x\rangle_{\mathbb{R}^n}\right)dx,$$

于是由 Parsaval 等式,

$$\begin{aligned}\sum_{j,k=1}^N\sum_{\alpha,\beta=1}^n\int_{\mathbb{R}^n}a_{\alpha\beta}^{jk}\partial_\alpha\varphi^j(x)\partial_\beta\varphi^k(x)dx&=4\pi^2\sum_{j,k=1}^N\sum_{\alpha,\beta=1}^n\int_{\mathbb{R}^n}a_{\alpha\beta}^{jk}\xi_\alpha\xi_\beta\hat\varphi^j(\xi)\hat\varphi^k(\xi)d\xi\\&\geqslant4\pi^2\sigma\int_{\mathbb{R}^n}|\xi|^2|\hat\varphi(\xi)^2d\xi\\&=\sigma\int_\Omega|\nabla\varphi(x)|^2dx.\end{aligned}$$

对于变系数情形, 可以利用单位分解, 在每个小邻域内把系数冻结为常数, 化归前面的估计; 所得的余项用 Schwarz 不等式合并到右边第二个积分中去. 因为写起来比较繁琐又离本课程的主题较远, 所以略去. (有兴趣的读者可以参看 K. Yosida, *Functional Analysis*, pp.175-177.) □

引理 6.3 设 $L \in C^2$ 满足严格 Legendre-Hadamard 条件, 即 $\exists \sigma > 0$, 使得

$$L_{p^i_\alpha p^j_\beta}(x,u,p)\xi^i\xi^j\eta_\alpha\eta_\beta \geqslant \sigma|\xi|^2|\eta|^2, \quad \forall (x,u,p) \in \Omega \times \mathbb{R}^N \times \mathbb{R}^{nN}.$$

又设存在 $\mu > 0$, 使得

$$Q_{u_0}(\varphi) \geqslant \mu \int_\Omega |\varphi|^2 dx, \quad \forall \varphi \in C_0^1(\Omega, \mathbb{R}^N),$$

则存在 $\lambda > 0$ 使得

$$Q_{u_0}(\varphi) \geqslant \lambda \int_\Omega (|\nabla\varphi|^2 + |\varphi|^2) dx, \quad \forall \varphi \in C_0^1(\Omega, \mathbb{R}^N).$$

从而 u_0 是 I 的一个严格弱极小点.

证明 因为 L 满足严格 Legendre-Hadamard 条件, 所以由 Gårding 不等式, 存在 $\alpha > 0, C_0 > 0$, 使得

$$\int_\Omega A_{u_0}(\nabla\varphi \cdot \nabla\varphi)dx \geqslant \int_\Omega [\alpha|\nabla\varphi|^2 - C_0|\varphi|^2]dx.$$

从

$$Q_{u_0}(\varphi) = \int_\Omega [A_{u_0}(\nabla\varphi\cdot\nabla\varphi) + 2B_{u_0}(\nabla\varphi\cdot\varphi) + C_{u_0}(\varphi\cdot\varphi)]dx,$$

推出有正常数 C_1, C_2, 使得

$$\begin{aligned}
&\alpha \int_\Omega |\nabla\varphi|^2 dx \\
&\leqslant Q_{u_0}(\varphi) + C_1\left(\left(\int_\Omega |\nabla\varphi|^2dx\right)^{1/2}\left(\int_\Omega |\varphi|^2dx\right)^{1/2} + \int_\Omega|\varphi|^2dx\right) + C_0\int_\Omega|\varphi|^2dx \\
&\leqslant \frac{\alpha}{2}\int_\Omega |\nabla\varphi|^2dt + Q_{u_0}(\varphi) + C_2\int_\Omega|\varphi|^2dx.
\end{aligned}$$

再应用

$$\int_\Omega |\varphi|^2 dx \leqslant \mu^{-1}Q_{u_0}(\varphi),$$

得到

$$\int_\Omega |\nabla\varphi(x)|^2dx \leqslant \frac{2}{\alpha}(1 + C_2\mu^{-1})Q_{u_0}(\varphi).$$

联合这两个不等式, 存在 $\lambda > 0$ 使得

$$Q_{u_0}(\varphi) \geqslant \lambda \int_\Omega (|\nabla \varphi|^2 + |\varphi|^2) dx.$$

根据定理 6.2, u_0 是 I 的一个严格弱极小点. □

我们再从 "特征值" 的角度给出严格弱极小点的一个判断.

设 $u_0 \in C^1(\bar{\Omega}, \mathbb{R}^N)$ 是 I 的 E-L 方程的解. 我们把

$$\boxed{\lambda_1 = \inf\{Q_{u_0}(\varphi) \mid \varphi \in C_0^1(\Omega, \mathbb{R}^N), \int_\Omega |\varphi(x)|^2 dx = 1\}}$$

称为 Jacobi 算子的第一特征值 (详见第十二讲).

定理 6.3 设 $L \in C^2$ 满足严格 Legendre-Hadamard 条件, 又设 $u_0 \in M$ 是 I 的一个弱极小点, 则 $\lambda_1 \geqslant 0$. 如果 $\lambda_1 > 0$, 那么 u_0 是 I 的一个严格弱极小点.

证明 事实上, 若 $\lambda_1 < 0$, 则 $\exists \varphi_0 \in C_0^1(\Omega, \mathbb{R}^N) \backslash \{\theta\}$, 使得

$$Q_{u_0}(\varphi_0) < \frac{\lambda_1}{2} \int_\Omega |\varphi_0(x)|^2 dx < 0.$$

这与 (6.6) 矛盾.

若 $\lambda_1 > 0$, 则

$$Q_{u_0}(\varphi) \geqslant \frac{\lambda_1}{2} \int_\Omega |\varphi(x)|^2 dx, \quad \forall \varphi \in C_0^1(\Omega, \mathbb{R}^N).$$

由引理 6.3 直接推出 u 是严格极小点. □

注 6.3 (强极小) 对于多重积分的变分问题也可以定义 Weierstrass 过度函数 $\mathfrak{E}_L \in C^1(\Omega \times \mathbb{R}^N \times \mathbb{R}^{nN} \times \mathbb{R}^{nN}, \mathbb{R}^1)$ 来描写的强极小的必要条件. 令

$$\mathfrak{E}_L(x, u, p, q) = L(x, u, q) - L(x, u, p) - \sum_{\alpha=1}^{n} \sum_{i=1}^{N} (q_\alpha^i - p_\alpha^i) L_{p_\alpha^i}(x, u, p). \quad \square$$

我们有

定理 6.3 设 $u_0 \in C^1(\bar{\Omega}, \mathbb{R}^N)$ 是 I 的一个强极小点, 则

$$\mathfrak{E}_L(x, u_0(x), \nabla u_0(x), \nabla u(x) + \pi) \geqslant 0, \ \forall x \in \Omega, \forall \pi = (\pi_\alpha^i), \ \ \mathrm{rank}(\pi) = 1.$$

证明的思想和一元情形类似, 但麻烦得多, 本书从略.

当 $n > 1$ 但 $N = 1$ 时, 也有类似的强极小的充分条件 (Lichtenstein 定理), 参看 [GH] p.390.

习 题

1. 设 $I(u) = \int_\Omega \sqrt{1+u_x^2+u_y^2}dxdy$, $\forall \varphi \in C_0^\infty(\Omega)$, 分别求第一与第二变分 $\delta I(u,\varphi), \delta^2 I(u,\varphi)$.

2. 设 $g = (g_{\alpha\beta}(x))_{1\leqslant\alpha,\beta\leqslant n}$ 是有界闭区域 Ω 上的一个连续正定矩阵函数, 记 $\det(g)$ 为其行列式, $(g^{\alpha\beta}(x))$ 为其逆矩阵. 令

$$I(u) = \frac{1}{2}\int_\Omega \sum_{\alpha,\beta=1}^n g^{\alpha\beta}(x)\partial_\alpha u(x)\partial_\beta u(x)\det(g)dx_1\cdots dx_n.$$

(1) 求它的 E-L 方程.

(2) 设 $\psi \in C^1(\partial\Omega,\mathbb{R}^1)$, $u_0 \in M := \{v \in C^1(\Omega,\mathbb{R}^1) \mid v|_{\partial\Omega} = \psi\}$ 是 I 在 M 上的一个临界点, 求证 u_0 是一个弱极小点.

(3) 求 J_{u_0}.

3. 求泛函

$$I(u) = \int_\Omega [|\nabla u|^p - |u|^q]dx \quad (1 \leqslant p,q < \infty)$$

的 E-L 方程. 设 u_0 是一个临界点, 求 J_{u_0}.

4. $u \in C^1(\mathbb{R}^4)$ 满足波动方程 $\Box u = 0$. 问 u 是不是泛函

$$I(u) = \int_{\mathbb{R}^4} [(\partial_t u)^2 - |\nabla_x u|^2]dtdx_1dx_2dx_3$$

的临界点或极小点?

5. 设 $F \in C^2(\mathbb{R}^1,\mathbb{R}^1)$ 满足 $|F''(t)| < \lambda_1$, Ω 上零边值 Laplace 算子的第一特征值

$$I(u) = \int_\Omega \left[\frac{1}{2}|\nabla u|^2 + F(u)\right]dx.$$

设 $u_0 \in C_0^1(\Omega)$ 是 I 一个临界点, 求证它是一个极小点.

6. 设 $\Omega \subset \mathbb{R}^2$ 是一个平面区域, ν 是一个常数, $f \in L^1(\Omega)$, $M = \{w \in C^2(\Omega,\mathbb{R}^1) | w|_\Omega = \partial_\nu w|_\Omega = 0\}$, 其中 ∂_ν 表示法向导数. 写出下列泛函的 E-L 方程:

$$I(w) = \int_\Omega [(w_{xx}+w_{yy})^2 - 2\nu(w_{xx}w_{yy} - w_{xy}^2) + fw]dxdy.$$

第七讲　约束极值问题

函数极值问题包括无条件极值问题和 (有) 条件极值问题；在泛函极值问题中也有无条件极值问题和 (有) 条件极值问题. 不过在变分问题中, 约束的形式更加多样化了.

§7.1　等周问题

所谓等周问题是指: 给定一个目标泛函 $(\mathcal{M}, I)$，一个约束泛函 $(\mathcal{M}, N)$ 以及一个给定的常数 c, 我们要求在条件 $N(u) = c$ 约束下, 求泛函 I 达到极小值的必要条件和充分条件, 即

$$\min\{I(u) \mid u \in \mathcal{M},\ N(u) = c\}.$$

下面这个例子是等周问题名称的由来.

例 7.1 (等周问题)　在平面上给定封闭曲线的弧长, 问什么样的曲线围成的面积最大?

我们用参数方程来描写一条封闭曲线:

$$\begin{cases} x = x(\theta), \\ y = y(\theta), \end{cases}$$

其中 $0 \leqslant \theta \leqslant 2\pi$ 是参数. 此曲线围成的面积是

$$A = \frac{1}{2}\int_0^{2\pi} (xy' - yx')d\theta,$$

它的长度是

$$L = \int_0^{2\pi} \sqrt{x'^2 + y'^2} d\theta.$$

如果给定弧长为 l, 那么我们的问题就是在约束 $L = l$ 之下, 求曲线 $(x(\theta), y(\theta))$ 使泛函 A 达到极小值. □

我们先回顾一下在数学分析中函数的条件极值是怎样求解的. 设 $f, g \in C^1(\Omega, \mathbb{R}^1)$, $\Omega \subset \mathbb{R}^n$ 是一个开集. 设 $g^{-1}(1) \neq \varnothing$. 如果 $x_0 \in \Omega$ 在约束条件 $g(x_0) = 1$ 下使得函数 f 达到极小值:

$$f(x_0) = \min_{g^{-1}(1)} f(x),$$

那么我们可以用 Lagrange 乘子法, 把条件极值问题转化为无条件极值问题去处理.

也就是说, 存在一个 Lagrange 乘子 $\lambda \in \mathbb{R}^1$, 使得

$$\nabla f(x_0) + \lambda \nabla g(x_0) = 0.$$

Lagrange 乘子有明确的几何意义: 如果 $\mathcal{M} = g^{-1}(1)$ 是一个微分流形, 那么 $\nabla g(x)$ 与它在点 $(x, g(x))$ 的外法向平行. 我们把 f 限制在 M 上, 记作 $\tilde{f} = f|_M$. 它的微分

$$d\tilde{f}(x) = \nabla f(x) - \frac{\nabla f(x) \cdot \nabla g(x)}{\|\nabla g(x)\|^2} \nabla g(x).$$

在极值点上应满足

$$d\tilde{f}(x) = 0 \iff \nabla f(x) + \lambda \nabla g(x) = 0,$$

其中

$$\lambda = -\frac{\nabla f(x) \cdot \nabla g(x)}{\|\nabla g(x)\|^2}$$

是 $-\nabla f(x)$ 在单位外法向量 $\nabla g(x)$(当 $\|\nabla g(x)\| = 1$ 时) 上的投影.

这表明引入 Lagrange 乘子 λ 以后, 条件极值问题的解是调整后的函数 $f + \lambda g$ 的稳定点 (临界点).

泛函的条件极值问题也可以通过 Lagrange 乘子法将其转化为另一个泛函的无条件极值问题. 我们有下列

定理 7.1 给定 $L, G \in C^2(\bar{\Omega} \times \mathbb{R}^N \times \mathbb{R}^{nN})$, $\rho \in C^1(\partial\Omega, \mathbb{R}^N)$, $\mathcal{M} = \{u \in C^1(\bar{\Omega}, \mathbb{R}^N) \,|\, u|_{\partial\Omega} = \rho\}$. 定义 $\mathcal{M}$ 上的泛函

$$I(u) = \int_\Omega L(x, u(x), \nabla u(x)) dx$$

与

$$N(u)=\int_\Omega G(x,u(x),\nabla u(x))dx.$$

设 c 是一个常数使得 $N^{-1}(c)\cap\mathcal{M}\neq\varnothing$. 若 $u_0\in\mathcal{M}$ 是 I 在约束 $N(u)=c$ 下的弱极小点, 即

$$I(u_0)=\min_{u\in\mathcal{M}\cap N^{-1}(c)} I(u),$$

又若 $\exists\varphi_0\in C_0^1(\Omega,\mathbb{R}^N)$, 使得 $\delta N(u,\varphi_0)\neq 0$, 则 $\exists\lambda\in\mathbb{R}^1$ 满足

$$\delta I(u_0,\varphi)+\lambda\delta N(u_0,\varphi)=0,\quad \forall\varphi\in C_0^1(\Omega,\mathbb{R}^N),$$

即, 若令

$$Q=I+\lambda N$$

为调整后的 Lagrange 函数, 则 u_0 满足对应于 Q 的 E-L 方程

$$\boxed{\begin{aligned}\sum_{\alpha=1}^n\partial_\alpha Q_{p_\alpha^i}(x,u_0(x),\nabla u_0(x))=Q_{u^i}(x,u_0(x),\nabla u_0(x)),\\ i=1,\cdots,N.\end{aligned}}\tag{7.1}$$

证明 注意到 $\varphi\mapsto\delta N(u_0,\varphi)$ 是线性的, 不妨设 $\delta N(u_0,\varphi_0)=1$.

我们在函数空间上把 $N^{-1}(c)\cap\mathcal{M}$ 看成一个超曲面. 如今对于与 φ_0 线性无关的任意 $\varphi\in C_0^1(\Omega,\mathbb{R}^N)$, 考虑通过点 u_0, 向量 φ_0 与向量 φ 所张成的平面 (见图 7.1)

$$\pi=\{u_0+\varepsilon\varphi+\tau\varphi_0\,|\,(\varepsilon,\tau)\in\mathbb{R}^2\},$$

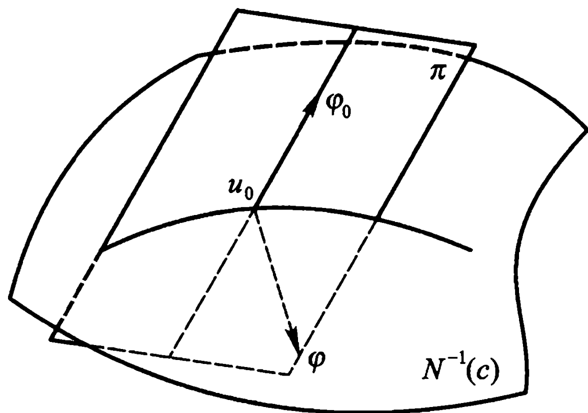


图 7.1

以及在 π 上 I 和 N 分别导出的两个函数

$$\Phi(\varepsilon,\tau)=I(u_0+\varepsilon\varphi+\tau\varphi_0)$$

和

$$\Psi(\varepsilon,\tau)=N(u_0+\varepsilon\varphi+\tau\varphi_0).$$

注意到

$$\Psi(0,0)=N(u_0)=c,\ \Psi_\tau(0,0)=\delta N(u_0,\varphi_0)=1,$$

当 $\varepsilon_0,\tau_0>0$ 充分小时, 我们就可以在 $R=(-\varepsilon_0,\varepsilon_0)\times(-\tau_0,\tau_0)$ 上应用隐函数定理: $\exists\xi\in C^1(-\varepsilon_0,\varepsilon_0)$, 使得 $(\varepsilon,\xi(\varepsilon))\in R$ 是方程

$$\Psi(\varepsilon,\tau)=c$$

在 R 中的唯一解. 这表明

$$\Psi(\varepsilon,\xi(\varepsilon))=c,\ \xi(0)=0,\ \xi'(0)=-\Psi_\varepsilon(0,0)=-\delta N(u_0,\varphi),$$

从而在小的 R 内,

$$N^{-1}(c)\cap\pi=u_0+\varepsilon\varphi+\xi(\varepsilon)\varphi_0.$$

令

$$g(\varepsilon)=I(u_0+\varepsilon\varphi+\xi(\varepsilon)\varphi_0).$$

因为 $v=u_0+\varepsilon\varphi+\tau\varphi_0\in\mathcal{M}$,

$$\|v-u_0\|_{C^1}\leqslant|\varepsilon|\|\varphi\|_{C^1}+|\tau|\|\varphi_0\|_{C^1},$$

而 u_0 是泛函 I 在 $M\cap N^{-1}(c)$ 上的弱极小点, 所以当 ε_0,τ_0 足够小时,

$$\Phi(\varepsilon,\xi(\varepsilon))\geqslant\Phi(0,0),$$

即, 0 是 g 的极小点.

由此推出

$$\begin{aligned}0&=g'(0)\\&=\frac{d}{d\varepsilon}\Phi(\varepsilon,\xi(\varepsilon))|_{\varepsilon=0}\\&=\Phi_\varepsilon(0,0)+\Phi_\tau(0,0)\xi'(0)\\&=\delta I(u_0,\varphi)+\lambda\delta N(u_0,\varphi),\end{aligned}\tag{7.2}$$

其中

$$\lambda = -\Phi_\tau(0,0)$$

与 φ 无关. 若令 $Q = I + \lambda N$, (7.2) 就是泛函 Q 的 E-L 方程

$$\mathrm{div} Q_p(x, u_0(x), \nabla u_0(x)) = Q_u(x, u_0(x), \nabla u_0(x)).$$ □

注 7.2 我们也把 λ 称为 Lagrange 乘子. □

注 7.3 可以考虑多个约束的泛函极值问题: 给定 $L, G^1, \cdots, G^m \in C^2(\bar{\Omega} \times \mathbb{R}^N \times \mathbb{R}^{nN})$, 函数集合 $\mathcal{M} \subset C^1(\bar{\Omega}, \mathbb{R}^N)$, 以及实数 $c_1, c_2, \cdots, c_m$. 令

$$I(u) = \int_\Omega L(x, u(x), \nabla u(x))dx,$$

$$N_j(u) = \int_\Omega G^j(x, u(x), \nabla u(x))dx, \quad j = 1, 2, \cdots, m,$$

如果 $u \in \mathcal{M}$ 是泛函 I 在约束

$$\int_\Omega G^j(x, u(x), \nabla u(x))dx = c_j \quad (j = 1, 2, \cdots, m)$$

下的极小点, 又如果 $\exists \varphi_k \in C_0^1(\Omega, \mathbb{R}^N), k = 1, 2, \cdots m$ 使得 $\det(a_{jk}) \neq 0$, 其中

$$\begin{aligned} a_{jk} &= \delta N_j(u, \varphi_k) \\ &= \int_\Omega [\langle G_u^j(x, u(x), \nabla u(x)), \varphi_k(x)\rangle + \langle G_p^j(x, u(x), \nabla u(x)), \nabla\varphi_k(x)\rangle]dx \\ &\quad (j, k = 1, \cdots, m), \end{aligned}$$

那么便有 Lagrange 乘子 $\lambda_1, \lambda_2, \cdots, \lambda_m$, 使得 u 满足对应于调节后的 Lagrange 函数

$$Q = I + \lambda_1 N_1 + \lambda_2 N_2 + \cdots + \lambda_m N_m$$

的 E-L 方程. □

例 7.1 (等周问题续) 由定理 7.1, 我们引入 Lagrange 乘子, 考虑调整后泛函的 E-L 方程

$$I(x, y) = A(x, y) + \lambda L(x, y),$$

得到

$$\begin{cases} -x' = \lambda \left(\dfrac{y'}{\sqrt{x'^2 + y'^2}} \right)', \\ y' = \lambda \left(\dfrac{x'}{\sqrt{x'^2 + y'^2}} \right)'. \end{cases}$$

解出来得到

$$\begin{cases} x - c_1 = -\lambda \dfrac{y'}{\sqrt{x'^2 + y'^2}}, \\ y - c_2 = \lambda \dfrac{x'}{\sqrt{x'^2 + y'^2}}, \end{cases}$$

其中 c_1, c_2 是常数. 这是圆的方程:

$$(x - c_1)^2 + (y - c_2)^2 = \lambda^2,$$

其半径 $r = \lambda = \frac{l}{2\pi}$, 而圆心是任意给定的点 (c_1, c_2). □

例 7.2 (特征值问题续) 在第一讲中我们曾经提到过如下的条件极值问题: 给定区域 $\Omega \subset \mathbb{R}^n$ 上的一个有界连续函数 $q \in C(\bar{\Omega})$, 定义泛函

$$I(u) = \int_\Omega |\nabla u|^2 dx, \quad N(u) = \int_\Omega q(x)|u(x)|^2 dx, \quad \mathcal{M} = C_0^1(\Omega).$$

求

$$\min\{I(u) \mid u \in \mathcal{M}, N(u) = 1\}.$$

按定理 7.1, 引入 Lagrange 乘子 λ, 调整后泛函 $Q = p^2 - \lambda q(x)u^2$ 对应的 E-L 方程是

$$-\Delta u = \lambda q u.$$

这里的 λ 正是 Laplace 算子 $-\Delta$ 对应于权函数 q 的特征值.

如果 $u_0 \in \mathcal{M}$ 是满足约束 $N(u) = \int_\Omega q(x)|u(x)|^2 dx = 1$ 的一个极小点, 那么 $\lambda = \int_\Omega |\nabla u_0(x)|^2 dx \neq 0$, 使得

$$-\Delta u_0(x) = \lambda q(x) u_0(x).$$

这个极小点 u_0 是 Laplace 算子 Δ 对于权函数 q 的一个特征函数, 而 Lagrange 乘子 λ 正是相应的特征值.

事实上, 引入 Lagrange 乘子 λ, 调整后的 Lagrange 函数导出的泛函

$$Q(u) = \int_\Omega [|\nabla u(x)|^2 - \lambda q(x) u^2(x)] dx$$

的所有临界点都是特征函数. □

§7.2 逐点约束

等周问题中的约束是积分形式的约束, 还有一类约束是逐点的.

例如, 给定一个函数 $M \in C^1(\bar{\Omega} \times \mathbb{R}^N \times \mathbb{R}^{nN}, \mathbb{R}^1)$, 我们要在 $\mathcal{M}$ 中, 并且在逐点约束

$$M(x, u(x), \nabla u(x)) = 0 \quad (\forall x \in \Omega)$$

的条件下, 求泛函

$$I(u) = \int_\Omega L(x, u(x), \nabla u(x)) dx$$

的极值.

大家自然要问, 现在还有没有类似的 Lagrange 乘子法可用? 下面我们只对完整 (holonomic) 约束的情形, 即当 M 仅依赖于 u (不依赖于 p) 时, 讨论这个问题.

定理 7.2 设 $\bar{\Omega} \subset \mathbb{R}^n$ 是一个有界闭集. 设 $L \in C^2(\bar{\Omega} \times \mathbb{R}^N \times \mathbb{R}^{nN}, \mathbb{R}^1)$, $M \in C^2(\mathbb{R}^N, \mathbb{R}^1), \rho \in C^1(\partial\Omega, \mathbb{R}^N)$, $\mathcal{M} = \{u \in PWC^1(\bar{\Omega}, \mathbb{R}^N) \,|\, u|_{\partial\Omega} = \rho\}$. 又设 $u_0 \in \mathcal{M}$ 是在上述约束下的一个局部极小点, 并且在有限个 $(n-1)$ 维逐片 C^1 的 $(n-1)$ 维超曲面之外是 C^2 的. 如果 $\forall x \in \bar{\Omega}, \nabla M(u_0(x)) \neq 0$, 那么存在一个连续函数 $\lambda \in C(\bar{\Omega})$, 使得 u_0 满足对应于调整后的 Lagrange 函数 $Q = L + \lambda M$ 的 E-L 方程:

$$\boxed{L_{u^i} + \lambda M_{u^i} = \sum_{\alpha=1}^{n} \frac{\partial L_{p_\alpha^i}}{\partial x_\alpha}, \quad 1 \leqslant i \leqslant N.} \tag{7.3}$$

证明 我们要逐片地构造局部连续函数 λ, 然后把它们粘连起来, 成为一个在 $\bar{\Omega}$ 上整体定义的连续函数.

1. $\forall x_0 \in \Omega, \exists$ 球 $B_r(x_0) \subset \Omega$, 使得 $\nabla_x(M(u_0(x))) \neq 0, \forall x \in B_r(x_0)$. 因为 $\nabla_u M(u_0(x)) \cdot \nabla u_0(x) = \nabla_x(M(u_0(x)))$, 所以 $\nabla_u M(u) \neq 0, \forall u \in u_0(B_r(x_0))$, $\nabla u_0(x) \neq 0, \forall x \in B_r(x_0)$. 不失一般性, 可设 $M_{u^N}(u_0(x)) \neq 0, \forall x \in B_r(x_0)$. 如果使用符号 $\tilde{u} = (u^1, \cdots, u^{N-1})$, 那么便可以解出 $u^N = U(\tilde{u})$, 其中 U 是一个 C^2 函数 (见图 7.2).

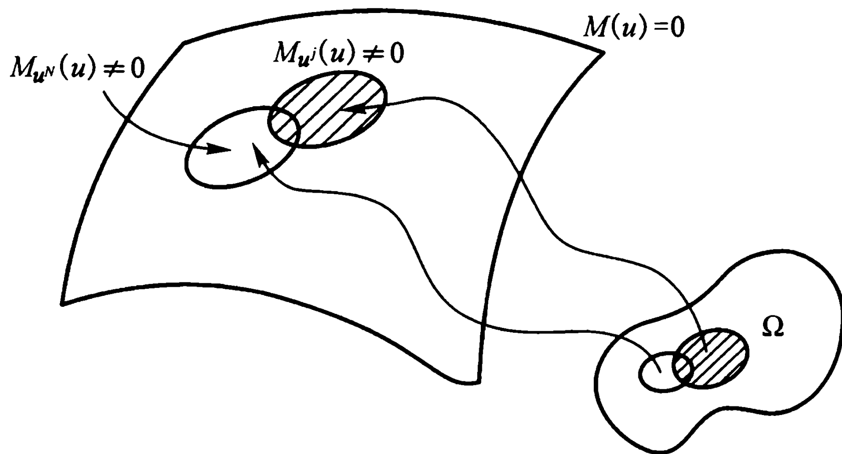


图 7.2

我们采用记号:

$$\tilde{p} = (p_\alpha^i)_{1\leqslant\alpha\leqslant n}^{1\leqslant i\leqslant N-1},$$

$$p^N = \left(\sum_{i=1}^{N-1} U_{u^i} p_\alpha^i\right)_{1\leqslant\alpha\leqslant n}$$

以及

$$P^N = \sum_{i=1}^{N-1} U_{u^i} \nabla u^i.$$

2. 因为 u_0 是一个极小点. 由局部化原理: 当把 I 的积分区域限制在 $B_r(x_0)$ 上时, $u_0|_{B_r(x_0)}$ 还是极小点. 如若不然, 则存在 $v \in PWC^1(\bar{B}_r(x_0))$, 使得 $v|_{\partial B_r(x_0)} = u_0|_{\partial B_r(x_0)}$, 并且

$$\int_{\bar{B}_r(x_0)} L(x, v(x), \nabla v(x))dx < \int_{\bar{B}_r(x_0)} L(x, u_0(x), \nabla u_0(x))dx,$$

那么我们在 $\bar{B}_r(x_0)$ 上用 $v(x)$ 替代 $u_0(x)$ 得到一个新的逐片可微函数, 其 I 值比 $I(u_0)$ 更小, 这与 u_0 是极小点矛盾.

令

$$\Lambda(x, \tilde{u}, \tilde{p}) = L(x, \tilde{u}, U(\tilde{u}), \tilde{p}, p^N),$$

因为 u_0 是在约束 $M(u) = 0$ 下 I 的极小点, 所以 $\tilde{u}_0$ 必是

$$J(\tilde{u}) = \int_{B_r(x_0)} \Lambda(x, \tilde{u}(x), \nabla \tilde{u}(x))dx$$

的极小点. 后者的 E-L 方程是

$$\Lambda_{u^i} = \mathrm{div}\Lambda_{p^i}(x, \tilde{u}(x), \nabla \tilde{u}(x)),\ i = 1, \cdots, N-1,$$

即

$$L_{u^i} + L_{u^N} U_{u^i} + \sum_{\alpha=1}^n L_{p_\alpha^N} \frac{\partial P_\alpha^N}{\partial u^i} = \sum_{\alpha=1}^n \frac{\partial}{\partial x_\alpha}(L_{p_\alpha^i} + L_{p_\alpha^N} U_{u^i}),\ i = 1, \cdots, N-1. \tag{7.4}$$

然而

$$\frac{\partial P_\alpha^N}{\partial u^i} = \sum_{j=1}^{N-1} \frac{\partial^2 U}{\partial u^i \partial u^j} u_\alpha^j = \frac{\partial U_{u^i}}{\partial x_\alpha}. \tag{7.5}$$

根据 (7.5), (7.4) 右端等于

$$\begin{aligned}&\sum_{\alpha=1}^{n}\left(\frac{\partial L_{p_\alpha^i}}{\partial x_\alpha}+U_{u^i}\frac{\partial L_{p_\alpha^N}}{\partial x_\alpha}+L_{p_\alpha^N}\frac{\partial U_{u^i}}{\partial x_\alpha}\right)\\=&\sum_{\alpha=1}^{n}\left(\frac{\partial L_{p_\alpha^i}}{\partial x_\alpha}+U_{u^i}\frac{\partial L_{p_\alpha^N}}{\partial x_\alpha}+L_{p_\alpha^N}\frac{\partial P_\alpha^N}{\partial u^i}\right).\end{aligned}$$

于是 (7.4) 化为

$$L_{u^i}+U_{u^i}\left(L_{u^N}-\sum_{\alpha=1}^{n}\frac{\partial L_{p_\alpha^N}}{\partial x_\alpha}\right)=\sum_{\alpha=1}^{n}\frac{\partial L_{p_\alpha^i}}{\partial x_\alpha},\tag{7.6}$$

注意到

$$M(\tilde{u_0},U(\tilde{u_0}))=0,$$

微分后有

$$M_{u^i}+M_{u^N}U_{u^i}=0,$$

即

$$U_{u^i}=-\frac{M_{u^i}}{M_{u^N}}.\tag{7.7}$$

在 $B_r(x_0)$ 上定义函数

$$\lambda=\frac{1}{M_{u^N}}(\mathrm{div}L_{p^N}-L_{u^N}).\tag{7.8}$$

将 (7.7) 和 (7.8) 代入 (7.6), 得到

$$L_{u^i}-\mathrm{div}L_{p^i}+\lambda M_{u^i}=0,\quad i=1,\cdots,N-1,\tag{7.9}$$

合并 (7.8) 与 (7.9) 就是在 $B_r(x_0)$ 上局部的 E-L 方程 (7.3).

3. 由于 $\bar{\Omega}$ 是紧的, 所以存在上述小球 $\{B_r(x_0)\,|\,x_0\in\Omega,\ r=r(x_0)\}$ 的一个有限开覆盖, 以及每个球上定义了的 $\lambda|_{B_r(x_0)}$. 设在 $B_{r_1}(x_1)$ 上 $M_{u^N}\neq 0$, 而在 $B_{r_2}(x_2)$ 上 $M_{u^j}\neq 0$. 我们有 $\lambda|_{B_{r_2}(x_2)}=\frac{1}{M_{u^j}}(\mathrm{div}L_{p^j}-L_{u^j})$, 而 $\lambda|_{B_{r_1}(x_1)}$ 就是 (7.8) 中的 λ, 那么当两个这样的小球有非空交时 ($B_{r_1}(x_1)\cap B_{r_2}(x_2)\neq\varnothing$), 由 (7.9) 即得

$$\lambda|_{B_{r_1}(x_1)}=\lambda|_{B_{r_2}(x_2)}.$$

我们把这些 $\lambda|_{B_r(x_0)}$ 粘接成一个整体的连续函数 λ. 这正是所要的结论. □

注 7.4 和等周问题一样, 我们也可以考虑不止一个约束函数的情形. 只要约束条件 $M_1(u) = M_2(u) = \cdots = M_s(u) = 0$ 满足

$$\operatorname{rank}\left(\frac{\partial M_i(u(x))}{\partial u^j}\right) = s, \quad \forall x \in \Omega.$$

这时调整后的 Lagrange 函数是

$$Q = I + \sum_{i=1}^{s} \lambda_i M_i,$$

其中 $\lambda_i \in C(\Omega),\ i = 1, \cdots, s.$ □

注 7.5 如果约束函数还依赖其他变量, 那么我们称之为非完整约束. 对于带非完整约束的变分问题还有没有 Lagrange 乘子可用? 这方面的情况比较复杂, 有兴趣的读者可以参看 Giaquinta, Hildebrandt 的书 [GH]. 其中的一种特殊情形, 称为最优控制问题, 我们将在第三部分介绍. □

例 7.3 (再看球面上的测地线) 我们曾经从无约束极值的角度讨论过球面上的测地线问题, 现在换一种看法, 把球面看成约束

$$x^2 + y^2 + z^2 = 1,$$

求泛函

$$I(u) = \int_a^b \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}dt$$

在这逐点约束条件下的极值.

令 $u = (x, y, z), p = (\xi, \eta, \zeta)$. 引入 Lagrange 乘子 $\lambda(t)$, 得到调整后的 Lagrange 函数

$$Q = \sqrt{\xi^2 + \eta^2 + \zeta^2} + \lambda(x^2 + y^2 + z^2).$$

它的 E-L 方程是

$$\frac{d}{dt}\left(\frac{\dot{u}}{|\dot{u}|}\right) = 2\lambda u,$$

还要满足 $\|u\| = 1$.

由于

$$\frac{d}{dt}\left(\frac{\dot{u}}{|\dot{u}|} \times u\right) = \frac{d}{dt}\left(\frac{\dot{u}}{|\dot{u}|}\right) \times u + \frac{\dot{u}}{|\dot{u}|} \times \frac{du}{dt} = 2\lambda u \times u + \frac{\dot{u}}{|\dot{u}|} \times \dot{u} = 0,$$

即得 $v = \frac{\dot{u}}{|\dot{u}|} \times u$ 是常向量.

这表明 $u \perp v$, 即测地线 u 必须位于垂直于常向量 v 的平面上, 因此这条曲线必是大圆的一部分. □

例 7.4 在斜面 $x=y$ 上, 求质量为 1 的质点受重力作用的运动方程. 把这质点的坐标记作 (x,y). Lagrange 函数是

$$L=\frac{1}{2}(p^2+q^2)-gy,$$

其中 g 是重力加速度, 约束是

$$M(x,y)=y-x=0.$$

调整后的 Lagrange 函数是

$$Q=L+\lambda M=\frac{1}{2}(p^2+q^2)-gy+\lambda(y-x).$$

相应的 E-L 方程是

$$\begin{cases}-\lambda=\dfrac{\partial Q}{\partial x}=\dfrac{dQ_p}{dt}=\ddot{x},\\ -g+\lambda=\dfrac{\partial Q}{\partial y}=\dfrac{dQ_q}{dt}=\ddot{y},\end{cases}$$

但因 $x=y$, 所以 $\lambda=g/2$. 解出得到:

$$x=y=-\frac{1}{2}\lambda t^2+\dot{x}(0)t+x(0),$$

其中 $\lambda=g/2$. □

例 7.5 (到球面的调和映射) 设 Ω 是 $\mathbb{R}^3$ 中的单位球, S^2 是 $\mathbb{R}^3$ 中的单位球面, $u=(u_1,u_2,u_3):\Omega\to S^2$. 如果 u 是下列约束极值问题的解,

$$\min\{I(u)\mid M(u)=0\},$$

其中

$$I(u)=\int_\Omega|\nabla u|^2dx=\int_\Omega\sum_{i=1}^{3}\sum_{\alpha=1}^{3}|\partial_\alpha u_i|^2dx,$$
$$M(u)=|u|^2-1=u_1^2+u_2^2+u_3^2-1,$$

那么我们把它称为是从单位实心球到单位球面的调和映射. 求这个调和映射 u 满足的微分方程.

解 写下调整后的 Lagrange 函数对应的 E-L 方程:

$$-\Delta u_i = \lambda u_i,$$

其中 Δ 是 Laplace 算子, 而 Lagrange 乘子 λ 是一个连续函数. 由

$$u_1^2 + u_2^2 + u_3^2 = 1,$$

求导得到

$$\langle u, \nabla u\rangle = 0,$$

其中 $\langle\cdot,\cdot\rangle$ 是 $\mathbb{R}^3$ 中的内积. 再求导又得到

$$\langle u, \Delta u\rangle + |\nabla u|^2 = 0.$$

在前面的 E-L 方程两边同时点乘以 u, 便有

$$\lambda = -\langle \Delta u, u\rangle = |\nabla u|^2,$$

即

$$-\Delta u = u|\nabla u|^2. \qquad \square$$

§7.3 变分不等式

除了等式约束外, 还有带不等式约束的变分问题.

例 7.6 (障碍问题) 设 $\Omega \subset \mathbb{R}^2$ 是平面上的一个有界开区域.
作为边界, 给定一个函数 $\varphi \in C^1(\partial\Omega)$.
作为障碍, 给定一个函数 $\psi \in C^1(\bar{\Omega})$.
作为外力, 给定一个函数 $f \in C(\bar{\Omega})$.
在 Ω 上, 我们考虑一张薄膜, 它的边界固定, $u|_{\partial\Omega} = \varphi$, 受外力作用, 但不得越过障碍, $u(x) \leqslant \psi(x), \forall x \in \bar{\Omega}$. $\square$

变分问题的提法如下: 求薄膜的平衡位置

$$u \in M = \{u \in PWC^1(\bar{\Omega}) \mid u|_{\partial\Omega} = \varphi\}.$$

在不等式约束下,

$$u(x) \leqslant \psi(x), \quad \forall x \in \Omega,\ u \in M,$$

使薄膜的能量 I 达到极小值:

$$I(u)=\int_\Omega\left[\frac{1}{2}|\nabla u(x)|^2-f(x)u(x)\right]dx.$$

一般地说, 作为定义域 M, 给定一个在 $\bar{\Omega}\subset\mathbb{R}^n$ 上定义的函数集合, 再给定 M 的一个凸子集 C, 以及一个 Lagrange 函数 L, 我们求

$$\min\left\{I(u)=\int_\Omega L(x,u(x),\nabla u(x))dx\,|\,u\in C\right\}.$$

设 $u\in C$ 是一个极小点. 我们要导出一个与 E-L 方程类似的关系. 事实上, 因为 C 是凸集, $\forall v\in C, tv+(1-t)u\in C$, $\forall t\in[0,1]$, 所以

$$I(tv+(1-t)u)\geqslant I(u),\quad \forall t\in[0,1].$$

由此推出

$$\delta I(u,v-u)=\lim_{t\to+0}[I(tv+(1-t)u)-I(u)]\geqslant 0,$$

即 $\forall v\in C$,

$$\boxed{\begin{aligned}&\int_\Omega[L_u(x,u(x),\nabla u(x))(v(x)-u(x))\\&+L_p(x,u(x),\nabla u(x))\cdot(\nabla v(x)-\nabla u(x))]dx\geqslant 0.\end{aligned}}$$

注意 E-L 方程与它的差别仅在于前者是一个等式, 而后者则是一个不等式. 所以我们把它称为变分不等式.

回到障碍问题, $C=\{u\in PWC^1(\bar{\Omega})\,|\,u(x)\leqslant\psi(x),\forall x\in\bar{\Omega},\ u|_{\partial\Omega}=\varphi\}$ 是一个凸集合, 所得的变分不等式是

$$\int_\Omega[\nabla u\nabla(v-u)-f(v-u)]dx\geqslant 0,\ \forall v\in C.$$

习 题

1. 求

$$\min\{I(u)\ |\,u\in C^1[0,1],\ u(0)=0,u(1)=2,N(u)=L\},$$

其中

$$I(u)=\int_0^1\dot{u}^2(t)dt,\ N(u)=\int_0^1 u(t)dt.$$

2. (Dido 问题) 求

$$\max\{I(u) \mid u \in C_0^1(0,b),\ N(u) = L\},$$

其中

$$I(u) = \int_0^b u(t)dt, \quad N(u) = \int_0^b (1 + \dot{u}(t)^2)^{1/2}dt.$$

3. 确定在等周约束 $\int_a^b r(t)u^2(t)dt = 1$ 下,

$$I(u) = \int_a^b [\ddot{u}^2(t) - p(t)\dot{u}^2(t) + q(t)u^2(t)]dt$$

的 E-L 方程, 其中 p, q, r 都是 $[a, b]$ 上的连续函数, u 满足边值条件: $u(a) = \dot{u}(a) = u(b) = \dot{u}(b) = 0$.

4. 求泛函

$$I(u) = \int_0^1 (\dot{u}_1^2 + \dot{u}_2^2)dt, \quad M = \{(u_1, u_2) \in C_0^1((0,1), \mathbb{R}^2)\}$$

在约束

$$u_2^2 + (u_1 - t) = 0$$

下的极值.

5. 设 $\Omega \subset \mathbb{R}^n$, $u : \Omega \to \mathbb{R}^N$, 又设 $G \in C^1(\mathbb{R}^N)$ 满足 $\nabla G(u) \neq 0,\ \forall\, u \in G^{-1}(1)$. 令

$$I(u) = \int_\Omega |\nabla u(x)|^2 dx.$$

求证约束变分问题

$$\min\{I(u) \mid G(u(x)) = 1,\ \forall x \in \Omega\}$$

的 E-L 方程是

$$-\Delta u^i = \lambda \partial_{u^i} G, \quad i = 1, \cdots, N,$$

$$\lambda = \frac{\sum_{k,j=1}^{N} \sum_{\alpha=1}^{n} \partial^2_{u^k u^j} G(u) \partial_\alpha u^k \partial_\alpha u^j}{|\nabla_u G(u)|^2}.$$

这是到超曲面 $G(u) = 1$ 的调和映射方程.

6. 设 $\Omega \subset \mathbb{R}^n$ 是一个有界开区域, $u : \bar{\Omega} \to \mathbb{R}^1$ 是曲面的非参数方程, 有给定的边值: $u|_{\partial\Omega} = \psi$ (ψ 是一个连续函数), 其面积为 $A(u) = \int_\Omega (1 + |\nabla u(x)|^2)^{1/2} dx$, 其下方体积为 $V(u) = \int_\Omega u(x)dx$. 求在给定下方体积为 V_0 的条件下, 使面积达到极小值的曲面应满足的方程.

7. 设 $X = (X_1(u,v), X_2(u,v), X_3(u,v))$, $\forall\, (u,v) \in B := \{(u,v) \,|\, u^2 + v^2 \leqslant 1\}$ 是曲面 S 的参数方程.

$$D(X) = \frac{1}{2} \int_B \sum_{i=1}^{3} |\nabla X^i|^2 dudv,$$

$$V(X) = \frac{1}{3} \int_B X \cdot (X_u \wedge X_v) dudv.$$

求在约束条件 $V(X) = V_0$ 下, 使 $D(X)$ 达到极小值的曲面 S 满足的方程.

第八讲　守恒律与 Noether 定理

物理学和力学中经常遇到各种守恒律, 例如能量守恒、动量守恒、角动量守恒等. E. Noether 注意到之所以会产生这些守恒律, 都是因为 Lagrange 函数具有某个群作用下的不变性.

§8.1　单参数微分同胚与 Noether 定理

1. 一个特殊的单参数函数族

给定一个有界区域 $\Omega \subset \mathbb{R}^n$, 一个函数 $u \in C^1(\bar{\Omega}, \mathbb{R}^N)$ 和一个 Lagrange 函数 $L \in C^2(\bar{\Omega} \times \mathbb{R}^N \times \mathbb{R}^{nN})$. 现在引入一族单参数微分同胚 $\eta_\varepsilon : \Omega \to \Omega_\varepsilon$, $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$, 其中 $\Omega_\varepsilon = \eta_\varepsilon(\Omega) \subset \mathbb{R}^n$ 是区域 Ω 的一族形变, 并且 $\eta_0 = id$ (图 8.1). 假设

$$\left.\frac{\partial \eta_\varepsilon(x)}{\partial \varepsilon}\right|_{\varepsilon=0} = \bar{X}(x).$$

那么形变:

$$y = \eta_\varepsilon(x) = x + \varepsilon \bar{X}(x) + o(\varepsilon).$$

再定义 $v : \mathbb{R}^n \times (-\varepsilon_0, \varepsilon_0) \to \mathbb{R}^N$ 满足: $v(x, 0) = u(x)$. 由此诱导出一族函数 $v_\varepsilon = v(\cdot, \varepsilon) : \Omega_\varepsilon \to \mathbb{R}^N, (y, \varepsilon) \in \Omega_\varepsilon \times (-\varepsilon_0, \varepsilon_0)$. 假设

$$\left.\frac{\partial v_\varepsilon(x)}{\partial \varepsilon}\right|_{\varepsilon=0} = \varphi(x).$$

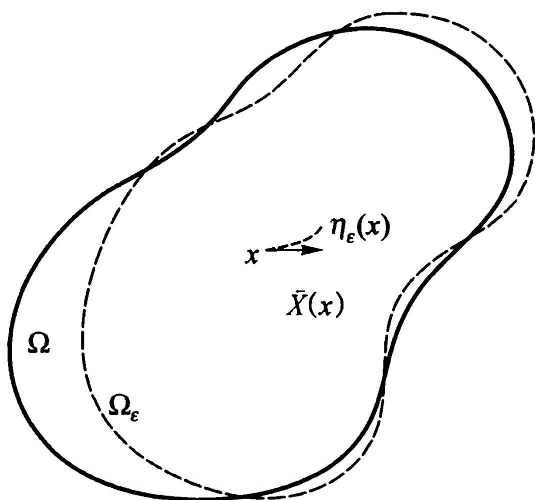


图 8.1

在前面我们讨论过的变分问题中, 在泛函的同一个定义域 M 里, 所有函数都有相同的定义区域 Ω, 其实这种限制并不是必需的, 可以让 M 中的函数有不同的定义域 Ω. 为了强调泛函对区域的依赖性, 我们记

$$I(u,\Omega)=\int_\Omega L(x,u(x),\nabla u(x))dx.$$

于是在这个函数族 $\{v_\varepsilon,\Omega_\varepsilon\}$ 上 I 的取值为

$$\begin{aligned}\Phi(\varepsilon)=I(v_\varepsilon,\Omega_\varepsilon)&=\int_{\Omega_\varepsilon} L(y,v_\varepsilon(y),\nabla v_\varepsilon(y))dy\\&=\int_\Omega L(\eta_\varepsilon(x),v_\varepsilon(\eta_\varepsilon(x)),\nabla_y v_\varepsilon(\eta_\varepsilon(x)))\det\left(\frac{\partial(\eta_\varepsilon(x))}{\partial x}\right)dx.\end{aligned}$$

因为

$$\frac{d}{d\varepsilon}\det\left(\frac{\partial(\eta_\varepsilon(x))}{\partial x}\right)\Bigg|_{\varepsilon=0}=\operatorname{div}(\bar{X}),$$

所以当 $u\in C^2$ 时, 有下列 Noether 恒等式:

$$\begin{aligned}\Phi'(0)&=\int_\Omega[\partial_{x_\alpha}L\bar{X}^\alpha+L_{u^i}\varphi^i+L_{p^i_\alpha}\varphi^i_{x^\alpha}+L\operatorname{div}\bar{X}]dx\\&=\int_\Omega[\operatorname{div}(L\bar{X})+L_{u^i}\varphi^i-\partial_\alpha(L_{p^i_\alpha})\varphi^i+\operatorname{div}(L_{p^i}\varphi^i)]dx,\end{aligned}$$

即

$$\Phi'(0)=\int_\Omega[E_L(u)\varphi+\operatorname{div}(L\bar{X}+L_{p^i}\varphi^i)]dx. \tag{8.1}$$

2. 一般的局部单参数变换群

现在我们在 "相空间"(即 x 与 u 同时变化的空间) 中作形变. 设 $\Omega\subset\mathbb{R}^n$. 又设 $\{\phi_\varepsilon\}:\Omega\times\mathbb{R}^N\to\mathbb{R}^n\times\mathbb{R}^N, |\varepsilon|<\varepsilon_0$ 为满足

$$\begin{cases} Y(x,u,0)=x, \\ W(x,u,0)=u \end{cases}$$

的一族映射 $(x,u)\mapsto(Y(x,u,\varepsilon),W(x,u,\varepsilon))$, 我们把它称为局部单参数变换群. 它有生成向量场

$$\boxed{\left.\frac{d\phi_\varepsilon}{d\varepsilon}\right|_{\varepsilon=0}=\sum_{\alpha=1}^{n}X^\alpha(x,u)\frac{\partial}{\partial x^\alpha}+\sum_{i=1}^{N}U^i(x,u)\frac{\partial}{\partial u^i},} \tag{8.2}$$

因此

$$\begin{cases} X(x,u)=\left.\dfrac{\partial}{\partial\varepsilon}Y(x,u,\varepsilon)\right|_{\varepsilon=0}, \\ U(x,u)=\left.\dfrac{\partial}{\partial\varepsilon}W(x,u,\varepsilon)\right|_{\varepsilon=0}. \end{cases}$$

如令 $\forall u\in C^1(\bar\Omega,\mathbb{R}^N)$, 我们要把它转化到前面所说的特殊的单参数函数族去. 令

$$\begin{cases} \eta(x,\varepsilon)=Y(x,u(x),\varepsilon), \\ \omega(x,\varepsilon)=W(x,u(x),\varepsilon), \end{cases}$$

则

$$\begin{cases} \eta(x,0)=x, \\ \omega(x,0)=u(x). \end{cases}$$

再令

$$\boxed{\begin{aligned} \bar X(x)&=\left.\frac{\partial\eta(x,\varepsilon)}{\partial\varepsilon}\right|_{\varepsilon=0}=X(x,u(x)), \\ \bar U(x)&=\left.\frac{\partial\omega(x,\varepsilon)}{\partial\varepsilon}\right|_{\varepsilon=0}=U(x,u(x)). \end{aligned}} \tag{8.3}$$

引入变量

$$y=\eta_\varepsilon(x)=\eta(x,\varepsilon)=x+\varepsilon\bar X(x)+o(\varepsilon)$$

以及区域的形变 $\Omega_\varepsilon=\eta_\varepsilon(\Omega)$, 则 $\eta_\varepsilon:\bar\Omega\to\bar\Omega_\varepsilon$ 是一个微分同胚.

它有逆映射 $\xi_\varepsilon=\eta_\varepsilon^{-1}$, 从而有

$$x=\xi_\varepsilon(y)=y-\varepsilon\bar X(y)+o(\varepsilon),\quad \forall y\in\Omega_\varepsilon.$$

又导出一族映射

$$v_\varepsilon(y) = v(y,\varepsilon) = \omega(\xi_\varepsilon(y),\varepsilon) = \omega(x,\varepsilon), \quad \forall y \in \Omega_\varepsilon.$$

令

$$\varphi(x) = \left.\frac{\partial v(x,\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0}, \quad \forall x \in \bar{\Omega},$$

推得

$$\bar{U} = \left.\frac{\partial v(\eta(x,\varepsilon),\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} = \varphi(x) + \sum_{\alpha=1}^{n} \partial_\alpha u(x)\bar{X}^\alpha(x),$$

即

$$\boxed{\varphi(x) = \bar{U} - \sum_{\alpha=1}^{n} u_{x_\alpha}\bar{X}^\alpha.} \tag{8.4}$$

这样一来, 对于给定的 $u \in C^2(\Omega,\mathbb{R}^N)$, 我们把具有生成向量场 (8.2) 的一般的局部单参数变换群 ϕ_ε 转化到一类特殊的单参数函数族, 其中 $\bar{X},\varphi$ 分别由 (8.3), (8.4) 确定. 代入 (8.1) 就得到在一般的局部单参数变换群下的 Noether 恒等式.

我们注意到 $I(u,\Omega)$ 在变换 $\{\phi_\varepsilon\}$ 下表示为

$$I(v_\varepsilon,\Omega_\varepsilon) = \int_{\Omega_\varepsilon} L(y, v_\varepsilon(y), \nabla v_\varepsilon(y))dy.$$

定义 8.1 $\forall \Omega_0 \subset \bar{\Omega}_0 \subset \Omega$, 令 $(\Omega_0)_\varepsilon = \eta(\Omega_0,\varepsilon)$. 如果

$$I(v_\varepsilon,(\Omega_0)_\varepsilon) = \text{const}\ (\text{不依赖于 } \varepsilon), \quad \forall\, u \in C^1(\bar{\Omega}_0,\mathbb{R}^N),$$

则 I 关于 $\{\phi_\varepsilon\}$ 是不变的.

事实上, 如果逐点有

$$L(\eta_\varepsilon(x), v_\varepsilon(\eta_\varepsilon(x)), \nabla_y v_\varepsilon(\eta_\varepsilon(x)))\det(\eta_\varepsilon(x)) = L(x,u(x),\nabla u(x)),$$

那么 I 便是 $\{\phi_\varepsilon\}$ 不变的.

总结前面的讨论我们得到

定理 8.1 (Noether) 设局部单参数变换群 $\{\phi_\varepsilon\}$ 是由向量场 (8.2) 生成的, 又设泛函

$$I(u) = \int_\Omega L(x,u(x),\nabla u(x))dx,$$

则 $\forall u \in C^2(\Omega, \mathbb{R}^N)$, 有 Noether 恒等式 (8.1), 其中

$$\begin{cases} \bar{X}(x) = X(x, u(x)), \\ \bar{U}(x) = U(x, u(x)), \\ \varphi^i(x) = \bar{U}^i(x) - \displaystyle\sum_{\alpha=1}^{n} \frac{\partial u^i}{\partial x^\alpha} \bar{X}^\alpha(x). \end{cases}$$

如果泛函 I 关于 $\{\phi_\varepsilon\}$ 是不变的, 那么

$$\boxed{E_L(u)\varphi + \operatorname{div}(L\bar{X} + L_{p^i}\varphi^i) = 0.}$$

在这里我们需要用 $(u(x), \nabla u(x))$ 代入 L, L_p 中的 (u, p).

进一步, 当 $u \in C^2$ 是 I 的弱极小点时, $(n-1)$ 形式

$$\nu = \sum_{\alpha=1}^{n} \left[L\bar{X}^\alpha + \sum_{i=1}^{N} L_{p^i_\alpha} \left(\bar{U}^i - \sum_{\beta=1}^{n} u^i_{x_\beta} \bar{X}^\beta \right) \right] dx^1 \wedge \cdots \wedge \hat{dx^\alpha} \wedge \cdots \wedge dx^n$$

是闭的, 即

$$d\nu = 0.$$

推论 8.1 当 $n = 1$ 时, 设 $\{\phi_\varepsilon\}$ 为 $\mathbb{R}^N$ 上的单参数局部变换群, 其生成向量场为 U (这意味着 $X = 0$), 若 I 关于 $\{\phi_\varepsilon\}$ 是不变的, 并且 $u \in C^2$ 是 I 的一个弱极小点, 则

$$\sum_{i=1}^{N} \bar{U}^i(u(t)) L_{p^i}(t, u(t), \dot{u}(t)) = \text{const}.$$

推论 8.2 设 $n = 1, L$ 是自守的 (即 L 与 t 无关), 则对于 I 的 E-L 方程的解 u 有

$$\left(L - \sum_{i=1}^{N} L_{p^i} p^i \right) \Bigg|_{(u,p)=(u(t),\dot{u}(t))} = \text{const}.$$

证明 事实上, 我们有单参数局部变换群 φ_ε, 其生成向量场为 $X = 1, U = 0$, 所以 $\varphi = \frac{du}{dt}$. □

§8.2 能动张量与 Noether 定理

当 $n = 1$ 时, Lagrange 函数 L 的 Legendre 变换是 Hamilton 函数 H. L 与 H 在公式中常处于对称的地位. 当 $n > 1$ 时, 因为 p 不再是向量而是一个张量, 所

以我们引进 Hamilton 能量动量张量 (简称能动张量) 来体现这种对称性:

$$T(x,u,p) = (T_\alpha^\beta(x,u,p)),$$

其中

$$T_\alpha^\beta = \sum_{i=1}^N p_\alpha^i L_{p_\beta^i} - \delta_\alpha^\beta L.$$

这就是说, T 的每个分量都是 L 的 Legendre 变换. 在这个意义上, Hamilton 能动张量是 Hamilton 函数在高维的推广.

例 8.1 设

$$L = \frac{1}{2}\left(\sum_{\beta,\gamma=0}^3 g_{\beta\gamma}p_\beta p_\gamma - M^2u^2\right) = \frac{1}{2}(p_0^2 - p_1^2 - p_2^2 - p_3^2 - M^2u^2),$$

其中 $g_{00}=1, g_{ii}=-1,\ i=1,2,3,$ 而 $g_{\beta\gamma}=0,\ \beta\neq\gamma$, 则

$$T_\beta^\alpha = \sum_{\gamma=0}^3 g_{\alpha\gamma}p_\beta p_\gamma - \delta_{\alpha\beta}L.$$

特别地,

$$T_0^0 = \frac{1}{2}(p_0^2+p_1^2+p_2^2+p_3^2+M^2u^2), \quad T_\beta^0 = p_0p_\beta, \quad \beta=1,2,3. \qquad \square$$

用 Hamilton 能动张量, Noether 定理最后又可以写成:

定理 8.2 设 $L\in C^2(\bar\Omega\times\mathbb{R}^N\times\mathbb{R}^{nN},\mathbb{R}^1)$. 又设

$$I(u) = \int_\Omega L(x,u(x),\nabla u(x))dx$$

关于局部单参数群 $\{\phi_\varepsilon\}$ 是不变的. 若 $u\in C^2(\bar\Omega,\mathbb{R}^N)$ 是 I 的一个弱极小点, 则

$$\boxed{\sum_{\alpha=1}^n\left(\sum_{i=1}^N L_{p_\alpha^i}\bar U^i - \sum_{\beta=1}^n T_\beta^\alpha\bar X^\beta\right)_{x_\alpha} = 0.}$$

简单记作

$$\operatorname{div}\left(L_{p^i}\bar U^i - T\bar X\right) = 0.$$

例 8.2 考虑 l 个质点 $m_1,\cdots,m_l$ 的质点系统: 位置坐标 $X=(X_1,\cdots,X_l)$, 其中 $X_i=(x_i,y_i,z_i)(1\leqslant i\leqslant l)$ 是第 i 个质点的空间坐标.

动能是

$$T=\frac{1}{2}\sum_{j=1}^{l}m_j\left|\dot{X}_j(t)\right|^2=\frac{1}{2}\sum_{j=1}^{l}m_j(\dot{x}_j^2+\dot{y}_j^2+\dot{z}_j^2),$$

位能是

$$V=-k\sum_{i<j}\frac{m_im_j}{|X_i-X_j|}=-k\sum_{i<j}\frac{m_im_j}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i-z_j)^2]^{1/2}}.$$

Lagrange 函数是

$$L=T-V.$$

对应的泛函是

$$I(X)=\int_{t_0}^{t_1}L(X(t),\dot{X}(t))dt. \qquad \square$$

• **空间平移群**$\{S_\varepsilon\}$.

设 $\{S_\varepsilon\}$ 是空间坐标依赖于参数 ε 的一族变换:

$$\widetilde{x}_i=x_i+\varepsilon,\quad \widetilde{y}_i=y_i,\quad \widetilde{z}_i=z_i,\quad 1\leqslant i\leqslant l.$$

因为 L 不依赖于 t, 所以 I 关于 $\{S_\varepsilon\}$ 是不变的. 对应的生成向量场 $X=0$, $U=(e_1,\cdots,e_1)$, $P_i^1=m_i\dot{x}_i(t)$, 其中 $e_1=(1,0,0)$. 此时由 Noether 定理得到

$$\sum_{i=1}^{l}P_i^1=\text{const}.$$

同理, 对 y,z 作平移, 分别有 $\sum_{i=1}^{l}P_i^2=\text{const},\sum_{i=1}^{l}P_i^3=\text{const}$. 记 $P=(P^1,P^2,P^3)$, 这就是

$$\sum_{i=1}^{l}P_i=\sum_{i=1}^{l}m_i\dot{X}_i(t)=\text{const},$$

即动量守恒.

• **时间平移群**.

设 $\{T_\varepsilon\}$ 是时空坐标依赖于参数 ε 的一族变换:

$$\widetilde{t}=t+\varepsilon,\quad \widetilde{x}_i=x_i,\quad \widetilde{y}_i=y_i,\quad \widetilde{z}_i=z_i,\quad 1\leqslant i\leqslant l,$$

则 I 关于 $\{T_\varepsilon\}$ 是不变的. 如令 $X=1,\ U=0$, 由 Noether 定理得到

$$H = pL_p - L = \sum_{j=1}^{l} m_j|\dot{X}_j|^2 - \left[\frac{1}{2}\sum_{j=1}^{l} m_j|\dot{X}_j^2| + k\sum_{i<j}\frac{m_i m_j}{[(x_i-x_j)^2+(y_i-y_j)^2+(z_i-z_j)^2]^{1/2}}\right] = T+V = \text{const}.$$

这是能量守恒.

- **单参数转动群**$\{R_\varepsilon\}$:

设 $\{R_\varepsilon\}$ 是时空坐标依赖于参数 ε 的一族变换:

$$\widetilde{t}=t,\quad \widetilde{x}_i = x_i\cos\varepsilon + y_i\sin\varepsilon,$$
$$\widetilde{y}_i = -x_i\sin\varepsilon + y_i\cos\varepsilon,\quad \widetilde{z}_i = z_i,\quad 1\leqslant i\leqslant l,$$

则 I 关于 $\{R_\varepsilon\}$ 是不变的. 如令

$$X=0,\ U=(Z_1,\cdots,Z_l),$$

其中 $Z_i=(y_i,-x_i,0), 1\leqslant i\leqslant l$, 所以

$$\sum_{i=1}^{l} m_i(y_i\dot{x}_i - x_i\dot{y}_i) = \sum_{i=1}^{l} L_{\dot{X}_i}Z_i = \text{const}.$$

对于平面 (y,z), (z,x) 作转动, 也有类似的等式, 这就是角动量守恒:

$$\sum_{i=1}^{l} m_i X_i\wedge\dot{X}_i = \text{const}.$$

例 8.3 考察一个物理场, 用 $u:\mathbb{R}^1\times\mathbb{R}^3\to\mathbb{R}^N$, $u=u(x)$ 表示场强分布, 其中 $x=(x_0,x_1,x_2,x_3), x_0=t$ 是时间坐标, $\bar{x}=(x_1,x_2,x_3)$ 是空间坐标. 给定一个 Lagrange 函数 $L:(\mathbb{R}^1\times\mathbb{R}^3)\times\mathbb{R}^N\times\mathbb{R}^{4N}$, 对应的泛函是

$$I(u)=\int_\Omega L(x,u(x),\nabla u(x))dx.$$

例如, $N=1$.

$$L(x,u,p)=\frac{1}{2}(p_0^2-p_1^2-p_2^2-p_3^2-M^2u^2).$$

这是 Klein-Gordon 场. □

根据狭义相对论, 任何物理场的 Lagrange 函数应该是在正向 Lorentz 变换下保持不变的. 正向 Lorentz 变换是保持二次型 $x_0^2 - x_1^2 - x_2^2 - x_3^2$ 不变的、保时间定向的四维时空线性变换. 一切正向 Lorentz 变换组成一个变换群, 称为正向 Lorentz 变换群.

平移变换群是正向 Lorentz 变换群的一个子群, 它有无穷小母元

$$\bar{X}^\beta = (\delta_\gamma^\beta)_{0\leqslant\gamma\leqslant3},$$

还有 $U = 0$, 所以由 Noether 定理得

$$\sum_{\alpha=0}^{3} \left(T_\gamma^\alpha\right)_{x_\alpha} = 0,\ \ \gamma = 0,1,2,3,$$

其中

$$T_\gamma^\alpha = \sum_\beta g_{\alpha\beta}\partial_\beta u \partial_\gamma u - \delta_{\alpha\gamma} L(x, u, \nabla u).$$

可以简洁地写成

$$\text{div} T_\gamma = 0,\ \ \gamma = 0,1,2,3,$$

其中 $T_\gamma = (T_\gamma^\alpha)_{0\leqslant\alpha\leqslant3}$.

任取 $[t_1, t_2] \times B_R(\theta) \subset \mathbb{R}^4$ 作为积分区域, 有

$$\int_{B_R(\theta)} T_\gamma^0(t_2, \bar{x}) d\bar{x} - \int_{B_R(\theta)} T_\gamma^0(t_1, \bar{x}) d\bar{x} + \int_{[t_1,t_2]\times\partial B_R(\theta)} T_\gamma^0(t, \bar{x}) \cdot \frac{\bar{x}}{|\bar{x}|} dt d\sigma = 0,$$

其中 $d\sigma$ 是二维球面的面积元, $|\bar{x}|$ 是 $\bar{x}$ 的模. 若当 $R \to \infty$ 时 $T_\gamma^0(t, x), |x| = R$, 一致地趋于零, 并且 $T_\gamma^0(t, \cdot)$ 在 $\mathbb{R}^3$ 上是可积的, 则

$$P_\gamma(t) = \int_{\mathbb{R}^3} T_\gamma^0(t, \bar{x}) d\bar{x} = \text{const}.$$

特别地, 取 $\gamma = 0$,

$$\begin{aligned} P_0(t) &= \int_{\mathbb{R}^3} T_0^0(t, \bar{x}) d\bar{x} \\ &= \int_{\mathbb{R}^3} \left[L_{p_0} u_{x_0} - L\right](t, \bar{x}) d\bar{x} \\ &= \int_{\mathbb{R}^3} \frac{1}{2}(|\partial_t u|^2 + |\nabla u|^2 + M^2 u^2) dx = \text{const}, \quad \forall t \in \mathbb{R}^1. \end{aligned}$$

这是能量守恒.

又取 $\gamma = 1, 2, 3$,

$$
\begin{aligned}
P_\gamma(t) &= \int_{\mathbb{R}^3} T_\gamma^0(t, \bar{x}) d\bar{x} \\
&= \int_{\mathbb{R}^3} [L_{p_0} u_{x_\gamma}](t, \bar{x}) d\bar{x} \\
&= \int_{\mathbb{R}^3} \partial_t u \partial_\gamma u = \text{const}, \quad \forall t \in \mathbb{R}^1.
\end{aligned}
$$

这是动量守恒.

除平移变换外, Lorentz 变换还有如下 6 个独立的转动母元.

$$
\varepsilon_{\mu\nu} = -\varepsilon_{\nu\mu}, \quad 0 \leqslant \mu \leqslant \nu \leqslant 3.
$$

作变换

$$
y_\mu = x_\mu + \sum_{\nu=0}^{3} g_{\nu\nu} \varepsilon_{\mu\nu} x_\nu,
$$

对应的向量场是

$$
\begin{aligned}
X_\mu^{\alpha\beta} &= \frac{dy_\mu}{d\varepsilon_{\alpha\beta}}\Big|_{\varepsilon_{\alpha\beta}=0} \\
&= g_{\beta\beta} x_\beta \delta_{\mu\alpha} - g_{\alpha\alpha} x_\alpha \delta_{\mu\beta}, \quad 0 \leqslant \beta < \alpha \leqslant 3.
\end{aligned}
$$

我们把

$$
M_{\alpha\beta\nu} = \sum_{\mu=0}^{3} (L_{p_\nu} u_{x_\mu} - \delta_{\mu\nu} L) X_\mu^{\alpha\beta} = \sum_{\mu=0}^{3} T_\mu^\nu X_\mu^{\alpha\beta}
$$

称为角动量. 根据 Noether 定理得

$$
\sum_{\mu,\nu=0}^{3} (T_\mu^\nu X_\mu^{\alpha\beta})_{x_\nu} = \sum_{\nu=0}^{3} \partial_\nu (g_{\beta\beta} T_\alpha^\nu x_\beta - g_{\alpha\alpha} T_\beta^\nu x_\alpha) = 0,
$$

推出

$$
\int_{\mathbb{R}^3} M_{\alpha\beta,0}(t) d\bar{x} = \text{const},
$$

$\forall (\alpha, \beta)$, $0 \leqslant \beta < \alpha \leqslant 3$. 这是角动量守恒.

注 8.1 对于电磁场、复标量场、Dirac 场等重要的物理场, 它们的守恒律都可以类似地由 Noether 定理推导出来. 由此可见, Noether 定理是一条非常基本的定理. □

§8.3 内极小

我们顺便再来讨论泛函取极小值的另一个必要条件. 以前我们对于一个泛函, 总是从因变量 u 变化导出泛函极值的必要条件: E-L 方程. 其实, 可以换一个角度, 把因变量 u 固定, 让自变量 x 变化. 因为这样也能得到不同的函数, 所以泛函也要跟着变化. 现在我们从这个方面来描写泛函取极小值应满足的条件.

沿用前面的记号, 设 $\eta_\varepsilon : \bar{\Omega} \to \bar{\Omega}$ 是一个微分自同胚,

$$y = \eta_\varepsilon(x) = x + \varepsilon \bar{X}(x) + o(\varepsilon),$$

其中 $\bar{X}|_{\partial\Omega} = 0$. η_ε 有逆变换

$$x = \xi_\varepsilon(y).$$

对于给定的函数 $u \in C^1(\bar{\Omega}, \mathbb{R}^N)$ 取

$$v_\varepsilon(y) = u(x),$$

即

$$v_\varepsilon = u \circ \xi_\varepsilon = u \circ \eta_\varepsilon^{-1}(y).$$

从而有

$$\begin{aligned} I(v_\varepsilon, \Omega) &= \int_\Omega L(y, v_\varepsilon(y), \nabla v_\varepsilon(y)) dy \\ &= \int_\Omega L\left(\eta_\varepsilon(x), u(x), \frac{\partial \xi_\varepsilon}{\partial y} \nabla u(x)\right) \det\left(\frac{\partial \eta_\varepsilon}{\partial x}\right) dx. \end{aligned}$$

注意到

$$\frac{\partial \xi_\varepsilon}{\partial y} = \left(\frac{\partial \eta_\varepsilon}{\partial x}\right)^{-1} = I - \varepsilon \frac{\partial \bar{X}}{\partial x} + o(\varepsilon),$$

所以当 $u \in C^2$, 有

$$\frac{d}{d\varepsilon} I(v_\varepsilon, \Omega)|_{\varepsilon=0} = \int_\Omega \sum_{\alpha=1}^n \left[-\sum_{\beta=1}^n \sum_{i=1}^N L_{p_\beta^i} \frac{\partial u^i}{\partial x_\alpha} \frac{\partial \bar{X}^\alpha}{\partial x_\beta} + L_{x_\alpha} \bar{X}^\alpha + L \frac{\partial \bar{X}^\alpha}{\partial x_\alpha}\right] dx$$

$$= \int_\Omega \left\{\sum_{\alpha=1}^n \left[L_{x_\alpha} - \partial_{x_\alpha}(L(x, u(x), \nabla u(x))) + \sum_{\beta=1}^n \partial_{x_\beta}\left(\sum_{i=1}^N L_{p_\beta^i} u_{x_\alpha}^i\right)\right] \bar{X}^\alpha\right\} dx$$

$$= \int_\Omega E_L(u) \cdot \left(-\sum_{\alpha=1}^n \frac{\partial u}{\partial x_\alpha} \bar{X}^\alpha\right) dx. \tag{8.5}$$

现在我们引入

定义 8.2 $u \in C^1(\bar{\Omega}, \mathbb{R}^N)$ 称为是 I 的一个内极小点, 如果 $\forall \bar{X} \in C_0^1(\Omega, \mathbb{R}^n)$, 泛函 I 在 v_ε 的变换下满足

$$\frac{d}{d\varepsilon} I(v_\varepsilon, \Omega)|_{\varepsilon=0} = 0.$$

于是当 $u \in C^2$ 是 I 的一个内极小点时, 按 (8.5), 我们有

$$\boxed{E_L(u) \frac{\partial u}{\partial x_\alpha} = 0.}$$

由此立得

推论 8.2 C^2 的弱极小点是其内极小点.

C^2 有内极小点的必要条件还可以通过能动张量来表达. $\forall u \in C^2(\bar{\Omega}, \mathbb{R}^N)$, $\forall X \in C_0^1(\Omega, \mathbb{R}^N)$,

$$\int_\Omega \sum_{\alpha=1}^n \left[\sum_{i=1}^N L_{p_\beta^i} \frac{\partial u^i}{\partial x_\alpha} \frac{\partial \bar{X}^\alpha}{\partial x_\beta} - L_{x_\alpha} \bar{X}^\alpha - L \frac{\partial \bar{X}^\alpha}{\partial x_\alpha} \right] dx$$
$$= -\int_\Omega \sum_{\alpha=1}^n \left[\sum_{\beta=1}^n \frac{\partial T_\alpha^\beta(x, u(x), \nabla u(x))}{\partial x_\beta} + L_{x_\alpha}(x, u(x), \nabla u(x)) \right] \bar{X}^\alpha dx,$$

所以有

$$\sum_{\beta=1}^n \frac{\partial T_\alpha^\beta(x, u(x), \nabla u(x))}{\partial x_\beta} + L_{x_\alpha}(x, u(x), \nabla u(x)) = 0, \quad \forall \alpha = 1, \cdots, n.$$

作为上述必要条件的一个应用, 我们有

例 8.4 (保形条件) 设 Ω 是一个平面区域. 又设 $u \in C^2(\bar{\Omega}, \mathbb{R}^2)$ 对于 $D(u) = \int_\Omega |\nabla u|^2 dxdy$ 是一个弱极小点, 则

$$\phi(z) := |\partial_x u|^2 - |\partial_y u|^2 + 2i \partial_x u \partial_y u$$

是 $z = x + iy$ $((x, y) \in \Omega)$ 的解析函数. □

证明 $\forall X = (X^1, X^2) \in C_0^1(\Omega, \mathbb{R}^2)$, 注意到 $L = \sum_{\alpha=1}^2 |p_\alpha|^2$ 不依赖于 z 和

u, 而 u 是一个内极小点, 按公式 (8.5) 的第一个等式,

$$\begin{aligned}&\int_\Omega\left[L\mathrm{div}X-\sum_{\alpha=1}^{2}\sum_{\beta=1}^{2}L_{p_\alpha}\partial_\beta u\partial_\alpha X^\beta\right]dxdy\\&=\int_\Omega[|\nabla u|^2(\partial_x X^1+\partial_y X^2)-\partial_x u(\partial_x u\partial_x X^1+\partial_y u\partial_x X^2)\\&\quad-\partial_y u(\partial_x u\partial_y X^1+\partial_y u\partial_y X^2)]dxdy\\&=\int_\Omega[(|\partial_x u|^2-|\partial_y u|^2)(\partial_x X^1-\partial_y X^2)+2\partial_x u\partial_y u(\partial_x X^2+\partial_y X^1)]dxdy\\&=0.\end{aligned}$$

令 $\xi=|\partial_x u|^2-|\partial_y u|^2$, $\eta=2\partial_x u\partial_y u$, 即

$$\int_\Omega[-(\partial_x\xi+\partial_y\eta)X^1+(\partial_y\xi-\partial_x\eta)X^2]dxdy=0.$$

因为 $X\in C_0^1(\Omega,\mathbb{R}^2)$ 是任意的, 所以

$$\begin{cases}\partial_x\xi+\partial_y\eta=0,\\\partial_y\xi-\partial_x\eta=0.\end{cases}$$

这是 Cauchy-Riemann 方程. 于是 ϕ 是解析的. □

§8.4 应用*

例 8.5 (Clairaut 定理) 设 l 是在光滑的旋转曲面 S 上的一条测地线, $\forall P\in l$, 用 $r(P)$ 表示在 P 点的水平圆的半径, $\alpha(P)$ 表示 l 与子午线在点 P 的夹角, 则

$$r(P)\sin\alpha(P)=\text{const}.$$

证明 用参数方程表示这旋转曲面:

$$(x,y,z)=(r\cos\theta,r\sin\theta,f(r)).\tag{8.6}$$

在它上面的一条曲线由 $r=r(\theta)$ 表示. 因为弧长泛函

$$L(r)=\int_{\theta_1}^{\theta_2}\sqrt{r^2+(1+f'(r)^2)\dot r^2}d\theta$$

不依赖于 θ, 所以有守恒律:

$$\frac{r^2}{\sqrt{r^2+(1+f'(r)^2)\dot r^2}}=\text{const}.$$

我们注意到, 弧长的微分是

$$ds = \sqrt{r^2 + (1 + f'(r)^2)\dot{r}^2}d\theta,$$

于是守恒律可以表示为

$$r^2\frac{d\theta}{ds} = \text{const.} \tag{8.7}$$

改用弧长 s 为参数: $r = r(s)$, $\theta = \theta(s)$, 曲线 l 的方程便是用 $r(s)$ 与 $\theta(s)$ 代入 (8.6) 中的 r 与 θ, l 的切向量就是

$$a = (\cos\theta, \sin\theta, f')\dot{r} + (-\sin\theta, \cos\theta, 0)r\dot{\theta}.$$

而水平的圆周方程是

$$(x, y, z) = (r\cos\theta, r\sin\theta, f(r)),$$

其中 $r = \text{const.}$ 它的切向量

$$b = (-r\sin\theta, r\cos\theta, 0).$$

这两个向量夹角的余弦是

$$a \cdot b = r(s)^2\dot{\theta}(s). \tag{8.8}$$ □

然而测地线与水平圆周的夹角 $\beta(s) = \frac{\pi}{2} - \alpha(s)$, 即

$$a \cdot b = r(s)\cos(\beta(s)) = r(s)\sin(\alpha(s)),$$

联合 (8.7) 与 (8.8), 我们得到

$$r(s)\sin(\alpha(s)) = \text{const.}$$ □

例 8.6 (Pohozaev 恒等式) 给定有光滑边界的区域 $\Omega \subset \mathbb{R}^3$, 以及 $g \in C(\mathbb{R}^1)$, 考察下列非线性椭圆型方程:

$$\begin{cases} -\Delta u = g(u) & \text{在 } \Omega \text{ 中}, \\ u = 0 & \text{在 } \partial\Omega \text{ 上}. \end{cases} \tag{8.9}$$

我们来证明当 $n \geqslant 3$ 时, 它的 (C^1 弱) 解满足下列恒等式:

$$\frac{n-2}{2}\int_\Omega |\nabla u|^2 dx - n\int_\Omega G(u)dx + \frac{1}{2}\int_{\partial\Omega}\left|\frac{\partial u}{\partial\nu}\right|^2 (x\cdot\nu)d\sigma = 0, \tag{8.10}$$

其中 G 是满足 $G(0) = 0$ 的 g 的原函数, $d\sigma$ 是 $\partial\Omega$ 上的面积元, 而 ν 是 $\partial\Omega$ 的单位外法向量.

证明 Lagrange 函数

$$L(u,p)=\frac{1}{2}p^2-G(u),$$

以及 $M=C_0^1(\Omega)$. 它们对应的 E-L 方程正是 (8.9).

不妨设原点 $\theta\in\Omega$, 设 $u\in C_0^1(\Omega)$, 考察 $\Omega_\varepsilon=(1+\varepsilon)\Omega$ 以及单参数微分同胚 $\eta_\varepsilon:\Omega\to\Omega_\varepsilon,\eta_\varepsilon(x)=(1+\varepsilon)x$ (见图 8.2). 令 $v_\varepsilon:\Omega_\varepsilon\to\mathbb{R}^1$,

$$v_\varepsilon(y)=u((1+\varepsilon)^{-1}y),$$

则 $\bar{X}=x$, 以及

$$\varphi(x)=\frac{d}{d\varepsilon}u((1+\varepsilon)^{-1}x)|_{\varepsilon=0}=-x\cdot\nabla u.$$

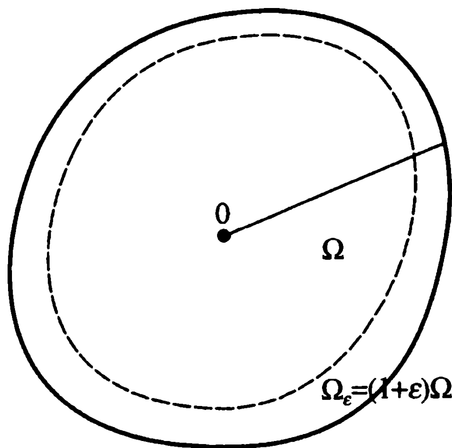


图 8.2

利用 Noether 恒等式, 我们知道

$$\begin{aligned}&\frac{d}{d\varepsilon}I(v_\varepsilon,\Omega_\varepsilon)|_{\varepsilon=0}\\&=\int_\Omega \operatorname{div}(L\bar{X}+L_p\varphi)dx\\&=\int_\Omega \operatorname{div}\left[\left(\frac{1}{2}|\nabla u|^2-G(u)\right)x-\nabla u(x\cdot\nabla u)\right]dx,\end{aligned}\tag{8.11}$$

一方面, 由 (8.9) 它等于

$$\begin{aligned}&\int_\Omega\left[\frac{n}{2}|\nabla u|^2-nG(u)+x\cdot\nabla\left(\frac{1}{2}|\nabla u|^2-G(u)\right)\right.\\&\left.-\Delta u(x\cdot\nabla u)-\nabla u\cdot\nabla(x\cdot\nabla u)\right]dx\\&=\int_\Omega\left[\frac{n}{2}|\nabla u|^2-nG(u)+\frac{x}{2}\cdot\nabla(|\nabla u|^2)-\frac{x}{2}\cdot\nabla(|\nabla u|^2)-|\nabla u|^2\right]dx\\&=\int_\Omega\left[\frac{n-2}{2}|\nabla u|^2-nG(u)\right]dx,\end{aligned}$$

另一方面对 (8.11) 应用 Green 公式, 注意到 $u|_{\partial\Omega}=0$, 得到

$$\frac{d}{d\varepsilon}I(v_\varepsilon,\Omega_\varepsilon)\bigg|_{\varepsilon=0}=-\frac{1}{2}\int_{\partial\Omega}\left|\frac{\partial u}{\partial\nu}\right|^2(x\cdot\nu)d\sigma.$$

即得

$$\frac{n-2}{2}\int_\Omega|\nabla u|^2dx-n\int_\Omega G(u)dx+\frac{1}{2}\int_{\partial\Omega}\left|\frac{\partial u}{\partial\nu}\right|^2(x\cdot\nu)d\sigma=0.$$

作为一个具体的应用, 我们取 $\Omega\subset\mathbb{R}^n$ $(n\geqslant 3)$ 为一个星形区域, 即, 从 θ 出发的任一射线与且仅与 $\partial\Omega$ 交于一点.

$$g(u)=u^{\frac{n+2}{n-2}}.$$

则方程

$$\begin{cases}-\Delta u=u^{\frac{n+2}{n-2}}, & x\in\Omega,\\ u|_{\partial\Omega}=0\end{cases}\tag{8.12}$$

没有非零解.

事实上, 我们取 $G(u)=\frac{n-2}{2n}u^{\frac{2n}{n-2}}$. 若 u 是 (8.12) 的一个解, 则由 (8.10),

$$\int_{\partial\Omega}\left|\frac{\partial u}{\partial\nu}\right|^2(x\cdot\nu)d\sigma=0.$$

因为 Ω 是一个星形区域, 所以 $x\cdot\nu>0$. 于是我们得到 $u|_{\partial\Omega}=\frac{\partial u}{\partial\nu}|_{\partial\Omega}=0$. 再由 Laplace 方程初值问题的唯一性, 得到 $u\equiv 0$. □

习 题

1. 设 $L(t,u,p)=t^2(p^2-\frac{1}{3}u^6)$. 又设 $\varphi_\varepsilon:\mathbb{R}^1\times\mathbb{R}^N\to\mathbb{R}^1\times\mathbb{R}^N$, 定义如下:

$$Y(t,u,\varepsilon)=(1+\varepsilon)t,\quad W(t,u,\varepsilon)=\frac{u}{(1+\varepsilon)^{1/2}}.$$

求证:

(1)

$$I(u)=\int_0^1 t^2\left(\dot{u}^2-\frac{1}{3}u^6\right)dt$$

是 φ_ε 不变的.

(2) 若 u 是 I 的 E-L 方程的解, 则

$$\frac{t^3}{3}u^6+t^3\dot{u}^2+t^2u\dot{u}=\text{const}.$$

2. 设 $L=(p+ku)^2$, 其中 k 是一个常数. 又设 $\{\varphi_\varepsilon\}$ 定义为

$$Y(t,u,\varepsilon)=t+\varepsilon,\quad W(t,u,\varepsilon)=u+\varepsilon\alpha e^{-kt},\quad \alpha\in\mathbb{R}^1.$$

(1) 验证: $I(u)=\int_0^1(\dot{u}+ku)^2dt$ 是 $\{\varphi_\varepsilon\}$ 不变的.

(2) 若 u 是 I 的 E-L 方程的解, 求 u 满足的守恒律.

(3) 用转化为 Hamilton 系统的方法解出 u.

(4) 对于这个 u, 验证 (2) 的结论.

第九讲　直接方法

§9.1　Dirichlet 原理与极小化方法

20 世纪之前的变分理论是以 E-L 方程为基础的, 和数学分析中求函数极值化归为求解临界点方程一样, 求解泛函极值也可化归为求解对应的 E-L 方程.

E-L 方程是微分方程. 当 $n=1$ 时, 这是常微分方程 (或常微分方程组), 只在一些特殊情形下, 才有可能求出它们的解析解; 当 $n>1$ 时, E-L 方程是偏微分方程, 能把一个偏微分方程的解析解写出来的情形更是少之又少!

19 世纪, 人们在研究电磁学和复变函数论中要求解 Laplace 方程

$$\begin{cases} \Delta u = 0 \quad \text{在 } \Omega \text{ 中}, \\ u|_{\partial\Omega} = \varphi \end{cases} \tag{9.1}$$

和 Poisson 方程

$$\begin{cases} \Delta u = f \quad \text{在 } \Omega \text{ 中}, \\ u|_{\partial\Omega} = \varphi. \end{cases} \tag{9.2}$$

Riemann 保形变换定理是一个著名的例子: 对于平面上的任意一个单连通区域 Ω, 存在一个保形变换, 使其一对一地映到单位圆盘. 为了证明这个保形变换的存在性, Riemann 把它化归为求解调和方程的边值问题 (9.1). 他注意到 (9.1) 是 Dirichlet 积分 (看成一个泛函)

$$D(u) = \int_\Omega |\nabla u(x)|^2 dx \tag{9.3}$$

在 $M = \{u \in C^1(\bar{\Omega}) \mid u \mid_{\partial\Omega} = \varphi\}$ 上的 E-L 方程, 因此可以通过求 D 的极小点来得到方程 (9.1) 的解.

这个思想开辟了求解偏微分方程的一条新途径. 如果一个微分方程是某个泛函的 E-L 方程, 那么我们可以反过来, 把求解偏微分方程的问题化归为求相应泛函的极值问题.

为了证明这个调和方程边值问题 (9.1) 有解, 只要证明 Dirichlet 积分在 M 上存在极小值. 然而为什么这个极小值是存在的? Riemann 使用了在 19 世纪中叶人们普遍相信的 Dirichlet 原理.

Dirichlet 原理 因为 D 有下界, 所以 "必有" 下确界, 即 "存在" u 使其达到极小值.

"理由" 如下: 取一列函数 $\{u_n\} \subset M$ 使得 $D(u_n) \to \inf_{u\in M} D(u)$. 因为 $\{u_n\}$ 是有界的, 所以有收敛的子列 $u_{n_k} \to u_0$. 这个 u_0 就是我们所要的解, $D(u_0) = \min_{u\in M} D(u)$.

对于这样的论证今天的读者显然是不能满意的. 就在 Riemann 保形变换定理证明之后不久, 人们展开了对 "Dirichlet 原理" 的讨论. 1870 年, Weierstrass 通过下例质疑上述的论证. 考虑下列极值问题:

$$I(u) = \int_{-1}^{1} x^2 u'^2 dx, \quad M = \{u \in C^1[-1,1] | u(-1) = -1,\ u(1) = 1\}.$$

(1) $\inf_{u\in M} I = 0$. 事实上, $I \geqslant 0$, 又取

$$u_\varepsilon = \frac{\arctan \dfrac{x}{\varepsilon}}{\arctan \dfrac{1}{\varepsilon}}, \quad \forall \varepsilon > 0,$$

则

$$I(u_\varepsilon) < \int_{-1}^{1} \frac{\varepsilon^2 (x^2+\varepsilon^2)^{-1}}{\left(\arctan \dfrac{1}{\varepsilon}\right)^2} dx = \frac{2\varepsilon}{\arctan \dfrac{1}{\varepsilon}} \to 0, \quad 当\ \varepsilon \to 0.$$

(2) 若 u_0 是 I 在 M 中的极小点, 则 $u_0' \equiv 0 \Rightarrow u_0 = \text{const}$. 这与边值条件矛盾!

受过数学分析严格化训练的读者当然理解: 并不是在任何情况下极小化序列都有子序列收敛到极小值.

给定一个拓扑空间 X, 假设函数 $f: X \to \mathbb{R}^1$ 下方有界, 即 $\exists M > 0$, 使 $f(x) > -M$, 从而有下确界 $m = \inf_{x\in X} f(x)$. 取极小化序列 $\{x_j\} \subset X$, 使 $f(x_j) \to m$. 在什么条件下 $\{x_j\}$ 有子列收敛到极小值?

记 $f_t = \{x \in X | f(x) \leqslant t\}, \forall t \in \mathbb{R}^1$. 如果

$$\exists t > m, \text{使得} f_t \text{ 是 “序列紧”的,} \tag{9.4}$$

那么当 N 充分大时, $\{x_j \,|\, j \geqslant N\} \subset f_t$. 于是有子列 $x_{j_k} \to x_0 \in f_t$.

至于这个 x_0 是不是极小点, 还应该再要求

$$f(x_0) \leqslant \underline{\lim} f(x_{j_k}).$$

我们把条件

$$x_n \to x_0 \Longrightarrow f(x_0) \leqslant \underline{\lim} f(x_n)$$

称为 f 的序列下半连续性 (有时在不会发生混淆时, 简作下半连续性).

总结起来, 若 $f: X \to \mathbb{R}^1$ 是序列下半连续的, 且 $\exists t > m$, $\exists$ 序列紧集 $K_t \supset f_t := \{x \in X | f(x) \leqslant t\}$, 则 f 在 X 上达到极小值.

下面我们要应用上面这个抽象定理解决变分问题.

在有限维欧氏空间, 对于下半连续函数 f, 只要添加强制性条件

$$f(x) \to +\infty \quad (\|x\| \to \infty), \tag{9.5}$$

就能推出 $\exists t > m$, 使得f_t 是 “序列紧” 的, 这是在数学分析课里大家都知道的事实.

但无穷维空间和有限维空间的情况大不一样. 例如在无穷维 Hilbert 空间 X 上, 考虑模的平方函数 $f(x) = \|x\|^2, \forall t > 0$, 它当然是强制的, 但集合

$$f_t = \{x | \|x\| \leqslant \sqrt{t}\}$$

在模拓扑下不是序列紧的！例如, $X = L^2[0, \pi]$, f_t 不是序列紧的, 它含有一个没有收敛子序列的序列 $\{\sqrt{2t/\pi} \sin nx\}_1^\infty$.

如果我们在 M 上取 C^1 拓扑, 那么一方面, 从 $\{D(u_n)\}$ 有界并不能推出 $\|u_n\|_{C^1}$ 有界; 另一方面, 也容易举例说明即使点列是 C^1 有界的, 也未必存在 C^1 收敛的子列. 这就是前面陈述的 “理由” 不能成立的原因.

§9.2 弱收敛与 *弱收敛

在有限维欧氏空间 $\mathbb{R}^m$, 我们常用 $x_n = (\xi_1^n, \cdots, \xi_m^n) \to x = (\xi_1, \cdots, \xi_m)$ 表示 x_n 收敛到 x. 这里收敛 “$\to$” 的意义可以按模收敛理解: $\|x_n - x\| = \left[\sum_{i=1}^m (x_i^n - x_i)^2\right]^{1/2} \to 0$, 也可以按坐标收敛理解: $x_i^n \to x_i$, $i = 1, \cdots, m$. 因为它们是等价的.

但在无穷维空间, 这二者则有很大区别! 例如在 l^2 空间, 给定 $x_n = (\xi_1^n, \xi_2^n, \cdots)$, $x = (\xi_1, \xi_2, \cdots)$, 其中 $\sum_{i=1}^{\infty} |\xi_i^n|^2 < \infty$, $\sum_{i=1}^{\infty} |\xi_i|^2 < \infty$. 我们说 x_n 按模收敛到 x (记作 $x_n \to x$) 是指

$$\|x_n - x\| = \left[\sum_{i=1}^{\infty}(x_i^n - x_i)^2\right]^{1/2} \to 0,$$

而 x_n 按坐标收敛到 x 是指

$$x_i^n \to x_i,\ i = 1, 2, \cdots.$$

若令 $\xi_i^n = \delta_{in}$, $n = 1, 2, \cdots$, 则 x_n 按坐标收敛到 0, 但不模收敛到任何元素.

如果在 l^2 空间中用 "坐标收敛" 来定义收敛, 那么有界点列确有收敛子列. 这和有限维空间的情形一样, 只要使用 "对角线法" 来论证就够了. 事实上, 设 $\{x_n\}$ 有界, 即 $\exists M > 0$, 使得 $\sum_{i=1}^{\infty}(\xi_i^n)^2 \leqslant M^2$. 由此推出 $|\xi_i^n| \leqslant M$, $\forall i,\ n \in \mathbb{N}$. 于是,

$$\begin{aligned}
&\exists\{n_k^1 \,|\, k \in \mathbb{N}\} \subset \mathbb{N} \text{ 使得 } \xi_1^{n_k^1} \to \xi_1^0,\\
&\exists\{n_k^2 \,|\, k \in \mathbb{N}\} \subset \{n_k^1 \,|\, k \in \mathbb{N}\} \text{ 使得 } \xi_2^{n_k^2} \to \xi_2^0,\\
&\cdots\\
&\exists\{n_k^l \,|\, k \in \mathbb{N}\} \subset \{n_k^{l-1} \,|\, k \in \mathbb{N}\} \text{ 使得 } \xi_l^{n_k^l} \to \xi_l^0,\\
&\cdots
\end{aligned}$$

当 $k \to \infty$ 时, 因为 $\forall N,\ \exists K = K(N)$, 当 $k > K$ 时,

$$\sum_{i=1}^{N} \left|\xi_i^0\right|^2 \leqslant \sum_{i=1}^{N} \left|\xi_i^{n_k^i}\right|^2 + 1 \leqslant M^2 + 1,$$

所以

$$\sum_{i=1}^{\infty} \left|\xi_i^0\right|^2 \leqslant M^2 + 1.$$

我们证明了对角序列 $x_{n_k^k}$ "坐标收敛" 到 $x_0 = (\xi_1^0, \xi_2^0, \cdots, \xi_k^0, \cdots) \in l^2$.

弱收敛与 *弱收敛的概念是由 "坐标收敛" 引发而来的.

定义 9.1 设 X 是一个线性赋范空间, 序列 $\{x_n\} \subset X$ 称为弱收敛到 x, 记作 $x_n \rightharpoonup x$, 如果对于任意 $x^* \in X^*$ 都有 $\langle x^*, x_n - x\rangle \to 0$, 其中 X^* 是 X 的共轭空间.

设 X^* 是一个线性赋范空间 X 的共轭空间, 序列 $\{x_n^*\} \subset X^*$ 称为 *弱收敛到 x^*, 记作 $x_n^* \rightharpoonup^* x^*$, 如果对于任意 $x \in X$ 都有 $\langle x_n^* - x^*, x\rangle \to 0$.

注 9.1 事实上, 在空间 X^* 既有弱收敛, 又有 *弱收敛的概念. 所谓弱收敛 $x_n^* \rightharpoonup x^*$, 是指对于任意 $x^{**} \in X^{**}$ 都有 $\langle x^{**}, x_n^* - x^* \rangle \to 0$; 所谓 *弱收敛, 是指对于任意 $x \in X$ 都有 $\langle x_n^* - x^*, x \rangle \to 0$. 因为有连续嵌入 $X \hookrightarrow X^{**}$, 所以弱收敛蕴含了 *弱收敛. □

显然, 模收敛蕴含了弱收敛以及 *弱收敛, 但反之不然.

例 9.1 在空间 $L^2(-\infty, \infty)$ 上, 任意给定一个具紧支集的非零函数 $\varphi(t)$, 令 $\varphi_n(t) = \varphi(t+n)$, 则 $\varphi_n \rightharpoonup 0$, 但 $\|\varphi_n\| = \|\varphi\| \neq 0$. □

例 9.2 注意到 $L^2[0, 2\pi]$ 是自共轭的, 它的共轭空间还是 $L^2[0, 2\pi]$. 考察序列 $\{\sin(nt)\} \subset L^2[0, 2\pi]$.

根据 Riemann-Lebesgue 引理, $\forall f \in L^1[0, 2\pi]$,

$$\int_0^{2\pi} f(t) \sin(nt) dt \to 0,$$

可以得到如下结论.

(1) 取 $X = L^p[0, 2\pi]$, $1 \leqslant p < \infty$, 把序列看成 $\{\sin(nt)\} \subset L^{p'}[0, 2\pi] = (L^p[0, 2\pi])^*$ 就有

$$\sin(nt) \rightharpoonup^* 0 \quad (n \to \infty).$$

(2) 取 $X = L^p[0, 2\pi]$, $1 \leqslant p < \infty$, 把序列看成 $\{\sin(nt)\} \subset L^\infty[0, 2\pi] \subset X$, $X^* = L^{p'}[0, 2\pi] \subset L^1[0, 2\pi]$, $1 < p' \leqslant \infty$ 就有

$$\sin(nt) \rightharpoonup 0 \quad (n \to \infty).$$

当 $1 < p < \infty$ 时, 空间是自反的, *弱收敛与弱收敛一致. □

更一般地, 我们有

例 9.3 设 $D = [0, 1]^N$ 是 $\mathbb{R}^N$ 上的单位立方砖. 设 $\varphi \in L^p(D)$, $1 \leqslant p \leqslant \infty$, 并将其作周期延拓, 令 $\varphi_n(x) = \varphi(nx)$, $\forall n \in \mathbb{N}$, 以及

$$\bar{\varphi} = \int_D \varphi(x) dx,$$

则

$$\varphi_n \rightharpoonup \bar{\varphi} \text{ 于 } L^p(D), \quad 1 \leqslant p < \infty,$$

以及

$$\varphi_n \rightharpoonup^* \bar{\varphi} \text{ 于 } L^\infty(D).$$

证明 首先, 不妨设 $\bar{\varphi}=0$. 因若不然, 则可用 $\tilde{\varphi}=\varphi-\bar{\varphi}$ 取代 φ.

其次, 我们注意到

$$\|\varphi_n\|_p^p=\int_D|\varphi(nx)|^p dx=\frac{1}{n^N}\int_{nD}|\varphi(y)|^p dy=\|\varphi\|_p^p,\quad \forall 1\leqslant p\leqslant\infty.$$

现在我们定义一个集合函数 $\Phi(E)=\int_E\varphi(x)dx$, 其中 E 是任意可测集. Φ 是 σ 可加的, 并且因为 φ 是周期的, 所以有 $\Phi(x+D)=0, \forall x\in\mathbb{R}^N$.

对于任意长方砖 $Q=\Pi_1^N(c_i,d_i)$, 在 nQ 内通过消去互不重叠的 D 的平移, 得到

$$\left|\int_D\chi_Q\varphi_n dx\right|=\left|\int_Q\varphi_n dx\right|=\frac{1}{n^N}|\Phi(nQ)|\leqslant\frac{N}{n}\int_D|\varphi|dx.$$

因此, 对于简单函数 $\xi=\sum_i\alpha_i\chi_{Q_i}$, $Q_i\cap Q_j=\varnothing, i\neq j$, 我们有

$$\int_D\varphi_n\xi dx\to 0,\quad n\to\infty.$$

对于 $1<p\leqslant\infty$, 上述简单函数组成 $L^{p'}(D)$ $\left(p'=\frac{p}{p-1}\right)$ 中的一个稠密集. 即 $\forall f\in L^{p'}(D), \exists\xi$ 为简单函数, 使得 $\|f-\xi\|_{p'}<\frac{\varepsilon}{2\|\varphi\|_p}$. 当 n 足够大时,

$$\left|\int_D\varphi_n f dx\right|\leqslant\|\varphi_n\|_p\|f-\xi\|_{p'}+\left|\int_D\varphi_n\xi dx\right|<\varepsilon.$$

对于 $p=1$ 的情形, 证明略作修改, 留作习题, 从略. □

§9.3 *弱列紧性

在变分学中使用弱收敛或 *弱收敛是希望有弱列紧性或 *弱列紧性. 我们把上一节讨论过的 l^2 空间有界序列有 "坐标收敛" 子列的结论和证明抽象化为以下定理.

定理 9.1 (Banach-Alaoglu) 设 X^* 是一个可分的线性赋范空间 X 的共轭空间. 设 $\{x_n^*\,|\,n=1,2,\cdots\}\subset X^*$ 是一个模有界的序列: $M=\sup\|x_n^*\|<\infty$, 则它必有一个 *弱收敛的子列.

证明 因为 X 是可分的, 所以有一个可数的稠密集合 $\{x_k\,|k=1,2,\cdots\}$.

对于 x_1, 因为 $|\langle x_n^*,x_1\rangle|$ 是有界的数列, 所以有子列 $x_{n_j^1}^*$, 使得 $\langle x_{n_j^1}^*, x_1\rangle$ 收敛.

对于 x_2, 因为 $|\langle x_{n_j^1}^*,x_2\rangle|$ 是有界的数列, 所以 $\{x_{n_j^1}^*\}$ 中还有子列 $\{x_{n_j^2}^*\}$, 使得 $\langle x_{n_j^2}^*, x_2\rangle$ 也收敛.

如此等等.

利用对角线法则, 可以抽出子列 $\{x^*_{n^j_j}\}$, 使得 $\langle x^*_{n^j_j}, x_k\rangle$ 收敛, $\forall k=1,2,\cdots$.

然而 $\{x_k\,|k=1,2,\cdots\}$ 是一个稠密集合, 而且 $\{x^*_n\,|\,n=1,2,\cdots\}\subset X^*$ 是模有界的, 所以 $\forall x\in X$, $\{\langle x^*_{n^j_j}, x\rangle\}$ 是收敛的序列.

定义

$$f(x)=\lim_{j\to\infty}\langle x^*_{n^j_j}, x\rangle.$$

易见 $f(x)$ 线性的, 并且它还是连续的,

$$|f(x)|\leqslant\sup_j\|x^*_{n^j_j}\|\|x\|\leqslant M\|x\|,\quad \forall x\in X.$$

从而 $\exists x^*\in X^*$, 使得 $f(x)=\langle x^*, x\rangle,\forall x\in X$, 即

$$x^*_{n^j_j}\rightharpoonup^* x^*.$$

□

于是我们得到如下基本定理.

定理 9.2 *设 X 是一个可分 Banach 空间的共轭空间 (例如, 自反的 Banach 空间). 又设 $E\subset X$ 是一个 *弱序列闭非空子集. 若 $f: E\longrightarrow\mathbb{R}^1$ 序列 *弱下半连续(简记为 s.w*.l.s.c), 并是强制的 (即 $\forall x\in E$, 当 $\|x\|\to\infty$ 时, $f(x)\to+\infty$), 则 f 在 E 上有极小值.*

证明 取 f 的一串极小化序列 $\{x_n\}\subset E$,

$$\lim f(x_n)=\inf_{x\in E}f(x).$$

由于 f 是强制的, $\{x_n\}$ 有界. 根据定理 9.1, $\{x_n\}$ 有 *弱收敛子列

$$x_{n_k}\rightharpoonup^* x_0.$$

由假设 E 是 *弱序列闭的, $x_0\in E$.

再利用 f 的序列 *弱下半连续性,

$$f(x_0)\leqslant\liminf f(x_{n_k}),$$

所以得到

$$f(x_0)=\inf_{x\in E}f(x).$$

□

现在回到 Dirichlet 积分. 一方面, 从 Dirichlet 积分 $D(u)$ 有界不能导出 C^1 模有界; 另一方面, 我们不知道 $C^1(\bar{\Omega})$ 是不是某个线性赋范空间的共轭空间. 由此可见, 要想验证 Dirichlet 原理, $C^1(\bar{\Omega})$ 不是一个适合的空间.

与 Dirichlet 积分 $D(u)$ 紧密联系的是如下 (半) 模:

$$\|u\| = \left(\int_\Omega |\nabla u|^2 dx\right)^{1/2},$$

以及模

$$\|u\|_1 = \left[\int_\Omega (|\nabla u|^2 + |u|^2) dx\right]^{1/2}$$

这模对应的内积是

$$(u, v) = \int_\Omega (\nabla u \cdot \nabla v + uv) dx.$$

但线性空间 $C^1(\bar{\Omega})$ 按这个模不是完备的, 我们把它的完备化空间记作 $H^1(\Omega)$. 这是一个 Hilbert 空间.

与半模 $D(u)$ 相关的 Dirichlet 内积是

$$D(u, v) = \int_\Omega \nabla u \cdot \nabla v dx,$$

二者之间的关系为 $D(u) = D(u, u)$.

然而当 Ω 是一个有界开集时, 在 $C_0^1(\bar{\Omega})$ 上我们可以把第三讲中一元函数的 Poincaré 不等式推广到多元情形.

引理 9.1 (Poincaré 不等式) 设 $\Omega \subset \mathbb{R}^n$ 是一个有界区域, $u \in C_0^1(\Omega)$, 则 $\exists C = C(\Omega)$ 使得

$$\int_\Omega |u|^2 dx \leqslant C \int_\Omega |\nabla u|^2 dx.$$

证明 取立方砖 $D \subset \mathbb{R}^n$ 使 $\Omega \subset D, \forall \varphi \in C_0^\infty(\Omega)$. 令

$$\tilde{\varphi}(x) = \begin{cases} \varphi(x), & x \in \Omega, \\ 0, & x \notin \bar{\Omega}, \end{cases}$$

则

$$\int_\Omega |\varphi|^2 dx = \int_D |\tilde{\varphi}|^2 dx, \quad \int_\Omega |\nabla \varphi|^2 dx = \int_D |\nabla \tilde{\varphi}|^2 dx.$$

记 $x = (x_1, \tilde{x})$. 由一元函数的 Poincaré 不等式,

$$\int_J |\tilde{\varphi}(x_1, \tilde{x})|^2 dx_1 \leqslant C \int_J |\partial_{x_1} \tilde{\varphi}(x_1, \tilde{x})|^2 dx_1,$$

其中 J 是 D 的沿 x_1 方向的投影. 再对 $\tilde{x}$ 积分即得

$$\int_D |\tilde{\varphi}|^2 dx \leqslant C \int_D |\nabla \tilde{\varphi}|^2 dx,$$

即

$$\int_\Omega |\varphi|^2 dx \leqslant C \int_\Omega |\nabla \varphi|^2 dx.$$

$\forall u \in C_0^1(\Omega)$, 通过取极限得到. □

这表明 $D(u)$ 是 $C_0^1(\Omega)$ 上的模. 我们把 $C_0^1(\Omega)$ 在 $H^1(\Omega)$ 中的闭包记作 $H_0^1(\Omega)$. 它是 Hilbert 空间 $H^1(\Omega)$ 中的一个闭子空间. $D(v)$, $D(v_0, v)$, $\forall\, v \in H_0^1(\Omega)$ 可以分别看成是 Dirichlet 积分与 Dirichlet 内积的连续扩张. 对给定的 $v_0 \in C^1(\bar{\Omega})$, 在 $M = v_0 + H_0^1(\Omega)$ 上, Dirichlet 积分定义为

$$D(u) = D(v_0) + 2D(v_0, v) + D(v),$$

其中 $u = v_0 + v \in M$.

如果我们把这样定义的 $D(u)$ 看成是 $H_0^1(\Omega)$ 上的泛函. 那么

1. $D(u)$ 是强制的.

我们要证明

$$D(v) \to \infty \Longrightarrow D(u) \to \infty.$$

从 Schwarz 不等式和 Young 不等式有

$$|D(v_0, v)| \leqslant D(v_0) + \frac{1}{4} D(v),$$

得到

$$D(u) \geqslant \frac{1}{2} D(v) - D(v_0).$$

强制性证毕.

2. $D(u)$ 是弱序列下半连续的.

设 $u_j = v_0 + v_j$, 则

$$u_j \rightharpoonup u(H^1(\Omega)) \Leftrightarrow v_j \rightharpoonup v(H_0^1(\Omega)).$$

注意到

$$|D(v_0, \varphi)| \leqslant D(v_0)^{1/2} D(\varphi)^{1/2}, \quad \forall\, \varphi \in H_0^1(\Omega),$$

$\varphi \mapsto D(v_0, \varphi)$ 是 $H_0^1(\Omega)$ 上的连续线性泛函. 因此有

$$D(v_0, v_j) \to D(v_0, v).$$

同理

$$D(v, v_j) \to D(v, v).$$

用 Schwarz 不等式得

$$D(v) = \lim D(v, v_j) \leqslant \liminf D(v)^{1/2} D(v_j)^{1/2},$$

即

$$D(v) \leqslant \liminf D(v_j).$$

这就是

$$D(u) \leqslant \liminf D(u_j).$$

$H_0^1(\Omega)$ 是一个 Hilbert 空间, 从而是自共轭的, 也可以直接验证 $H_0^1(\Omega)$ 是一个自共轭空间, $H_0^1(\Omega)$ 便是 Banach 空间 $H_0^1(\Omega)$ 的共轭空间. 空间 $H_0^1(\Omega)$ 还是可分的 (详细证明见下一讲). 现在我们利用定理 9.2 就推出 Dirichlet 积分在 $H_0^1(\Omega)$ 上达到极小值. 即我们验证了 Dirichlet 原理.

注 9.2 为了验证 Dirichlet 原理, 其实还有一个更为直接的方法 —— 正交投影法. 在几何上这相当于在一个超平面上求一点到一定点的距离达到极小值 (参看第十二讲). 由此导出 Riesz 表示定理以及 Hilbert 空间的自共轭性.

在这个验证过程中并不需要用到弱下半连续性以及弱列紧性, 只需要空间 $H_0^1(\Omega)$ 的完备性就够了. 但这个方法可应用的范围十分有限. □

注 9.3 在这个例子中, 我们得到的解 u_0 属于空间 $H^1(\Omega)$. 我们并不知道它是否可微, 更不知道它是否 C^2. 进一步的讨论必须明确: u_0 在什么意义下是调和方程的解? □

对于更一般的泛函 I 连同边值条件, 为了要想用极小化序列的方法证明解的存在性, 我们需要做到:

1. 建立合适的函数空间, 它是一个自反的 Banach 空间, 或是可分的 Banach 空间的共轭空间.

2. 泛函 I 对于这空间的拓扑是序列 * 弱下半连续的.

3. 泛函 I 对于这空间的拓扑是强制的.

4. 要明确所得到的极小解在什么意义下是给定方程的解 (满足弱的形式的 E-L 方程)?

5. 要进一步讨论所得到的极小解有没有满足方程的可微性?

§9.4 自反空间与 Eberlein-Schmulyan 定理*

在泛函分析中对于弱列紧性有深入的讨论, 可以利用空间的自反性假设避开空间的可分性要求, 导出弱列紧性 (这时 *弱列紧性与弱列紧性没有区别).

我们来回顾一下自反空间的定义. 一个 Banach 空间 X 的共轭空间 X^* 也是一个 Banach 空间, 所以还可以再考虑 X^* 的共轭空间 $(X^*)^*$, 记作 X^{**}, 并称其为 X 的第二共轭空间.

注意到 $\forall x \in X$, 可以在 X^* 上定义一个泛函

$$\langle F_x, x^* \rangle = \langle x^*, x \rangle, \quad \forall x^* \in X^*,$$

F_x 在 X^* 上是线性的, 满足

$$|\langle F_x, x^* \rangle| \leqslant \|x^*\| \|x\|.$$

从而 F_x 还是连续的, 满足

$$\|F_x\| \leqslant \|x\|.$$

我们称 $T : x \to F_x$ 为自然映射. 实际上, $T : X \to X^{**}$ 是一个连续嵌入. 应用 Hahn-Banach 定理, $\exists x^* \in X^*$, 使得 $\|x^*\| = 1$, $\langle x^*, x \rangle = \|x\|$, 便有

$$\|x\| = \langle x^*, x \rangle = \langle F_x, x^* \rangle \leqslant \|F_x\|.$$

我们得到, T 还是一个等距同构, 即 X 与它的第二共轭空间 X^{**} 的一个闭线性子空间是等距同构的.

定义 9.2 一个 Banach 空间称为是自反的, 如果以上定义的等距映射 T 是满射.

Banach 有一个关于可分性的结论.

定理 9.3 (Banach) 设 X 是一个线性赋范空间. 若它的共轭空间 X^* 是可分的, 则 X 本身也是可分的.

证明 1. 记 S_1^* 为 X^* 的单位球面, 则 S_1^* 是可分的. 事实上, 记 $\{x_n^*\} \subset X^*$ 为一个可数稠子集. 令 $y_n^* = x_n^* / \|x_n^*\|$, 则 $\{y_n^*\}$ 是 S_1^* 上的一个可数稠子集. 因为 $\forall x^* \in S_1^*, \exists x_{n_j}^*$, $\|x_{n_j}^* - x^*\| \to 0$, 从而, $\|x_{n_j}^*\| \to 1$, 且

$$\|y_{n_j}^* - x^*\| \leqslant \left| 1 - \frac{1}{\|x_{n_j}^*\|} \right| \|x_{n_j}^*\| + \|x_{n_j}^* - x^*\| \to 0.$$

所以 $\{y_n^*\}$ 是 S_1^* 的可数稠子集.

2. 由定义, 存在 $x_n \in X$, 使得 $\|x_n\| = 1$, $\langle y_n^*, x_n \rangle \geqslant 1/2$. 令 $X_0 = \overline{\text{span}\{x_n\}}$. 要证明 $X_0 = X$. 若不然, 由 Hahn-Banach 定理, 存在 $y^* \in S_1^*$, 使得 $\langle y^*, x \rangle = 0, \forall x \in X_0$. 但

$$\|y^* - y_n^*\| \geqslant |\langle y^* - y_n^*, x_n \rangle| \geqslant \frac{1}{2}.$$

这是不可能的. 所以 $X_0 = X$, 即得 X 是可分的. □

我们有以上两个定理的如下推论.

推论 9.1 设 X 是一个可分、自反的 Banach 空间, 则它的任意有界序列必有一个弱收敛子列.

其实, 在这个推论中可分性的假设还可以去掉, 这要用到如下定理.

定理 9.4 (Pettis) 自反 Banach 空间 X 的闭线性子空间 X_0 是自反的.

证明 我们要证明 $\forall x_0^{**} \in X_0^{**}$, $\exists x \in X_0$, 使得

$$\langle x_0^{**}, x_0^* \rangle = \langle x_0^*, x \rangle, \quad \forall x_0^* \in X_0^*. \tag{9.6}$$

1. 定义映射 $T: X^* \to X_0^*$ 为 $Tx^* = x^*|_{X_0}$, 即

$$\langle Tx^*, x_0 \rangle = \langle x^*|_{X_0}, x_0 \rangle, \quad \forall x_0 \in X_0.$$

T 是线性连续的. 它有共轭映射 $T^* \in L(X_0^{**}, X^{**})$, $\forall x_0^{**} \in X_0^{**}, T^* x_0^{**} \in X^{**}$.

2. 因为 X 是自反的, 所以 $\exists x \in X$, 使得

$$\langle x^*, x \rangle = \langle T^* x_0^{**}, x^* \rangle = \langle x_0^{**}, x^*|_{X_0} \rangle, \quad \forall x^* \in X^*. \tag{9.7}$$

我们要证明 $x \in X_0$. 用反证法, 若 $x \notin X_0$, 则对闭线性子空间 X_0 应用 Hahn-Banach 定理, $\exists x_1^* \in X^*$ 使得

$$\langle x_1^*, x \rangle \neq 0, \quad x_1^*|_{X_0} = 0.$$

但

$$\langle x_1^*, x \rangle = \langle x_0^{**}, x_1^*|_{X_0} \rangle = 0.$$

这是一个矛盾.

3. 证明 (9.6), 即 x_0^{**} 是 $x \in X_0$ 在自然映射下的对应.

根据 Hahn-Banach 定理, $\forall x_0^* \in X_0^*, \exists x^* \in X^*$ 使得 $x^*|_{X_0} = x_0^*$. 这时 (9.7) 蕴含了

$$\langle x_0^{**}, x_0^* \rangle = \langle x^*, x \rangle = \langle x_0^*, x \rangle, \quad \forall x_0^* \in X_0^*.$$

我们证明了 X_0 是自反的. □

推论 9.2 若 X 是一个自反的 Banach 空间, 则 X^* 是自反的.

证明 "$\Rightarrow$". $(X^*)^{**} = (X^{**})^* = X^*$.

"$\Leftarrow$". 若 X^* 自反, 则由 $\Rightarrow$, X^{**} 自反. 而 X 是 X^{**} 的一个闭线性子空间, 再由 Pettis 定理, X 也是自反的. □

于是有

定理 9.5 (Eberlein-Schmulyan) 自反 Banach 空间 X 中的任意有界序列 $\{x_n\}$ 必有一个弱收敛子列.

推论 9.3 Hilbert 空间中的任意有界序列 $\{x_n\}$ 必有一个弱收敛子列.

除了 Hilbert 空间之外, 我们再看一些自反空间的例子.

考虑 Lebesgue 函数空间 $L^p(\Omega)$, 其中 $\Omega \subset \mathbb{R}^n, 1 \leqslant p < \infty$, 我们知道

$$(L^p(\Omega))^* = L^{p'}(\Omega), \quad 1 \leqslant p < \infty, p^{-1} + p'^{-1} = 1,$$

它的含义是: 对于 $L^p(\Omega)$ 上的任意连续线性泛函 F 存在在 a.e. 意义下唯一的 $L^{p'}(\Omega)$ 函数 v, 使得它可以表示为

$$F(u) = \int_\Omega v(x)u(x)dx,$$

并且

$$\|F\|_{(L^p)^*} = \left(\int_\Omega |v(x)|^{p'} dx\right)^{1/p'} = \|v\|_{p'}.$$

对 $\forall\, u \in L^p(\Omega),\ 1 < p < \infty$,

$$v \mapsto \int_\Omega u(x)v(x)dx$$

可以看成是空间 $L^{p'}(\Omega)$ 上的连续线性泛函 G. 同理, 还存在 $w_u \in L^p(\Omega)$, 使得

$$G(v) = \int_\Omega w_u(x)v(x)dx, \quad \forall\, v \in L^{p'}(\Omega),$$

即

$$\int_\Omega u(x)v(x)dx = G(v) = \int_\Omega w_u(x)v(x)dx, \quad \forall v \in L^{p'}(\Omega).$$

所以, $w_u(x) = u(x)$, a.e..

按上述意义, 当 $1 < p < \infty$ 时,

$$(L^p(\Omega))^{**} = (L^{p'}(\Omega))^* = L^p(\Omega),$$

即空间 $L^p(\Omega)$ 是自反的.

习 题

1. 试证: 对于空间 l^2 上的有界点列, 坐标收敛 $\Leftrightarrow$ 弱收敛.

2. 给定一族函数

$$u_\varepsilon = \frac{\arctan \dfrac{x}{\varepsilon}}{\arctan \dfrac{1}{\varepsilon}}, \ \forall \varepsilon > 0.$$

问: 它们在空间 $L^2(-1,1)$ 上是否模收敛或弱收敛? 在空间 $H^1(-1,1)$ 上是否模收敛或弱收敛?

3. 在 $L^2([0,2\pi])$ 上, 求

$$w\text{-}\lim_{n\to\infty} \sin nt \text{ 和 } w\text{-}\lim_{n\to\infty} \sin^2 nt.$$

4. 在空间 $H^1(0,1)$ 上给定一个函数序列

$$u_j(t) = \begin{cases} t - \dfrac{k}{j}, & t \in \left[\dfrac{k}{j}, \dfrac{2k+1}{2j}\right], \\ -t + \dfrac{k+1}{j}, & t \in \left[\dfrac{2k+1}{2j}, \dfrac{k+1}{j}\right], \end{cases} \qquad k = 0, 1, \cdots, j-1,$$

$j = 1, 2, \cdots$. 它们是否模收敛或弱收敛?

第十讲 Sobolev 空间

我们曾指出 C^1 和 C_0^1 空间不是验证 Dirichlet 原理的合适空间, 取而代之的是 H^1 和 H_0^1.

在用直接方法求解变分问题时这种情况有普遍性, 因为泛函是含导数的变分积分, 而同阶导数的 C 类空间的 C 型模是由逐点的模量的极大值决定的, 而 C 型模是不可能被这个变分积分所控制的. 此外, 为了具备序列弱列紧性, 空间必须取成线性赋范空间的共轭空间, 这样的空间至少应该是完备的. 本讲介绍的 Sobolev 空间是满足这些要求的函数空间.

§10.1 广义导数

设 $u, v \in L^1_{\text{loc}}(\Omega)$, $\Omega \subset \mathbb{R}^n$, $\forall i = 1, 2, \cdots, n$, 称 v 是 u 的对 x_i 的广义偏导数 $v = D_{x_i}u$, 是指 $\forall \varphi \in C_0^\infty(\Omega)$,

$$\int_\Omega v\varphi dx = -\int_\Omega u\partial_{x_i}\varphi dx.$$

更一般地, 对于多重指标 $\alpha = (\alpha_1, \cdots, \alpha_n)$, 用记号:

$$|\alpha| = \sum_{i=1}^n \alpha_i, \quad \partial^\alpha = \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n}.$$

定义 10.1 称 v 是 u 的 α 阶广义导数, 如果 $\forall \varphi \in C_0^\infty(\Omega)$,

$$\int_\Omega v\varphi dx = (-1)^{|\alpha|}\int_\Omega u\partial^\alpha\varphi dx.$$

记作 $v = D^\alpha u$.

例 10.1 若 $n=1, J=(-1,1), u(x)=|x|$. 因为

$$\begin{aligned}\int_{-1}^{1}|x|\varphi'(x)dx &= \int_0^1 x\varphi'(x)dx + \int_{-1}^{0}(-x)\varphi'(x)dx\\ &= -\int_{-1}^{1}\text{sgn}x\varphi(x)dx,\end{aligned}$$

所以 $D(|x|)=\text{sgn}x$. □

例 10.2 若 $u \in C^k(\Omega)$, 则 $D^\alpha u = \partial^\alpha u, \forall\ \alpha,\ |\alpha| \leqslant k$. 这是因为

$$\int_\Omega \partial^\alpha u\varphi dx = (-1)^{|\alpha|}\int_\Omega u\partial^\alpha\varphi dx.$$ □

§10.2 空间 $W^{m,p}(\Omega)$

定义 10.2 ($W^{m,p}(\Omega)$ 空间) 设 $p \in [1,\infty], m \in \mathbb{N}$, 令

$$W^{m,p}(\Omega) := \{u \in L^p(\Omega) \mid D^\alpha u \in L^p(\Omega),\ \forall\,\alpha,\ |\alpha| \leqslant m\},$$

规定范数为

$$\begin{aligned}\|u\|_{m,p} &= \left(\sum_{|\alpha|\leqslant m}\int_\Omega |D^\alpha u(x)|^p dx\right)^{1/p},\quad 1 \leqslant p < \infty,\\ \|u\|_{m,\infty} &= \operatorname*{esssup}_{x\in\Omega}\sum_{|\alpha|\leqslant m}|D^\alpha u(x)|.\end{aligned}$$

我们称其为 Sobolev 空间.

显然 Sobolev 空间是线性赋范空间.

对于有界区域 Ω, 显然有

$$\begin{aligned}W^{m,\infty}(\Omega) \subset W^{m,q}(\Omega) \subset W^{m,p}(\Omega) \subset W^{m,1}(\Omega),&\quad 1<p<q<\infty,\\ W^{m,p}(\Omega) \subset W^{l,p}(\Omega),&\quad 0 \leqslant l \leqslant m,\end{aligned}$$

以及

$$\text{若 } \Omega_1 \subset \Omega_2,\ u \in W^{m,q}(\Omega_2),\ \text{则 } u|_{\Omega_1} \in W^{m,q}(\Omega_1).$$

定理 10.1 $W^{m,p}(\Omega)$ 空间是完备的, 从而是 Banach 空间.

证明 设 $\{u_j\}$ 是一个基本列, 则 $\forall\,\alpha, |\alpha| \leqslant m$, $\{D^\alpha u_j\}$ 都是 L^p 的基本列. 从而存在 $g_\alpha \in L^p(\Omega)$, 使得 $D^\alpha u_j \to g_\alpha$, $L^p(\Omega)$.

如今 $\forall\,\varphi \in C_0^\infty(\Omega)$, 我们有

$$\langle D^\alpha u_j, \varphi\rangle = (-1)^{|\alpha|}\langle u_j, \partial^\alpha \varphi\rangle,$$

即得

$$\langle D^\alpha u_j, \varphi\rangle \to \langle g_\alpha, \varphi\rangle, \quad \forall\,\alpha,\ |\alpha| \leqslant m,$$

以及

$$\langle u_j, \partial^\alpha \varphi\rangle \to \langle g_0, \partial^\alpha \varphi\rangle, \quad \forall\,\alpha,\ |\alpha| \leqslant m.$$

这就是

$$\langle g_\alpha, \varphi\rangle = (-1)^{|\alpha|}\langle g_0, \partial^\alpha \varphi\rangle.$$

由此推出

$$g_\alpha = D^\alpha g_0,$$

以及

$$\| D^\alpha u_j - g_\alpha \|_p \to 0, \quad \forall\,\alpha, |\alpha| \leqslant m.$$

所以, $g_0 \in W^{m,p}(\Omega)$, 并且有

$$\|u_j - g_0\|_{m,p} \to 0,\ \ j \to \infty.$$

□

我们指出: 虽然广义导数是对偶地、整体地定义的, 它与局部定义的普通导数的关系还是相当密切的. 特别地, 对于一维 Sobolev 空间的函数, 它们的广义导数就是几乎处处导数! 设 $n = 1, J = [a, b]$, 我们有

例 10.3 $W^{1,1}(J) = AC(J)$, J 上的绝对连续函数空间, 并且

$$Du(x) = u'(x), \quad \text{a.e.}. \tag{10.1}$$

证明 $\forall u \in W^{1,1}(J)$, 我们证明

$$u(x) - u(a) = \int_a^x Du(t)dt, \quad \forall\,x \in J.$$

$\forall n \in \mathbb{N}$, 令

$$\varphi_n(t) = \begin{cases} n(t-a), & t \in \left[a, a+\frac{1}{n}\right], \\ 1, & t \in \left[a+\frac{1}{n}, x-\frac{1}{n}\right], \\ -n(x-t), & t \in \left[x-\frac{1}{n}, x\right], \\ 0, & t \in [x, b]. \end{cases}$$

因为 $\exists \xi_{n,k} \in C_0^\infty(J), \|\xi_{n,k}\|_{C^1} \leqslant 2n$, 使得

$|\varphi_n(t) - \xi_{n,k}(t)| \rightrightarrows 0$ 在 J 上一致, $|\varphi_n'(t) - \xi_{n,k}'(t)| \to 0, k \to \infty$, a.e. $t \in J$.

所以由 $\int_J u(t)\xi_{n,k}'(t)dt = -\int_J Du(t)\xi_{n,k}(t)dt$ 导出

$$\int_J u(t)\varphi_n'(t)dt = -\int_J Du(t)\varphi_n(t)dt.$$

令 $n \to \infty$, 即得

$$u(x) - u(a) = \int_a^x Du(t)dt, \quad \forall x \in J.$$

由此可见 $u(x)$ 是绝对连续函数, 并且

$$u'(x) = Du(x), \quad \text{a.e. } x \in J.$$

反之, $\forall u \in AC(J), u'(x)$ 几乎处处存在, 并且属于 $L^1(J)$. 只需再证

$$Du(x) = u'(x), \quad \text{a.e. } x \in J.$$

事实上, 因为 $u(x) = \int_y^x u'(t)dt + u(y), \forall x, y \in J$, 所以

$$\int_J u(x)\varphi'(x)dx = -\int_J u'(x)\varphi(x)dx, \quad \forall \varphi \in C_0^\infty(J). \qquad \square$$

此即 (10.1)

例 10.4 $W^{1,\infty}(J) = \mathrm{Lip}(J)$, 后者是 J 上的 Lipschitz 函数空间.

证明 "$\supset$". 设 $u \in \mathrm{Lip}(J)$, 则 u 是绝对连续函数. 从而几乎处处有导数 u' 满足: $u(y) - u(x) = \int_x^y u'(t)dt$, 并且

$$|u'(x)| \leqslant \sup_{y \in J} \frac{|u(y) - u(x)|}{|y - x|} \leqslant M,$$

由例 10.3 得 $Du \in L^{\infty}$, 即 $u \in W^{1,\infty}$, 以及

$$\|u\|_{1,\infty} \leqslant \|u\|_{\text{Lip}}.$$

“$\subset$”. 反之, 若 $u \in W^{1,\infty}(J)$, 则由 (10.1),

$$|u(y) - u(x)| \leqslant \int_x^y |u'(t)|\, dt \leqslant \|Du\|_{\infty} |y - x|, \quad \forall x, y \in J.$$

我们得到

$$\|u\|_{\text{Lip}} \leqslant \|u\|_{W^{1,\infty}}. \qquad \square$$

定义 10.3 我们记 $C_0^{\infty}(\Omega)$ 在 $W^{m,p}(\Omega)$ 中的闭包为 $W_0^{m,p}(\Omega)$.

引理 10.1 设 $u \in W^{m,p}(\Omega)$, $\psi \in C_0^{\infty}(\Omega)$, 则 $(\psi u) \in W_0^{m,p}(\Omega)$, 而且 $\text{supp}(D^{\alpha}(\psi u)) \subset \text{supp}(\psi)$, $|\alpha| \leqslant m$.

证明 只证明 $m = 1$ 的情况, 其余用数学归纳法. 事实上, $\forall \varphi \in C_0^{\infty}(\Omega)$,

$$\begin{aligned}\int_{\Omega} D_{x_i}(\psi u)\varphi dx &= -\int_{\Omega} \psi u \partial_{x_i}\varphi dx \\ &= -\int_{\Omega} u[\partial_{x_i}(\psi\varphi) - \partial_{x_i}\psi\varphi]dx \\ &= \int_{\Omega} [\partial_{x_i}\psi u + \psi D_{x_i} u]\varphi dx.\end{aligned}$$

我们得到

$$D_{x_i}(\psi u) = \psi(D_{x_i} u) + (\partial_{x_i}\psi)u. \qquad \square$$

§10.3 泛函表示

为了了解 Banach 空间 $W^{m,p}(\Omega)$(或 $W_0^{m,p}(\Omega)$) 的弱拓扑, 我们需要考虑这些空间的泛函表示.

我们知道 $L^p(J)$ 的共轭空间是

$$(L^p(\Omega))^* = L^{p'}(\Omega), \quad 1 \leqslant p < \infty, \quad \frac{1}{p} + \frac{1}{p'} = 1,$$

即 $\forall f \in (L^p(\Omega))^*, \exists v \in L^{p'}(\Omega)$ 使得

$$\langle f, u \rangle = \int_{\Omega} u(x)v(x)dx, \quad \forall u \in L^p(\Omega),$$

并且

$$\|f\| = \sup_{\|u\|_p \leqslant 1} \langle f, u \rangle = \left(\int_\Omega |v|^{p'} \right)^{1/p'}$$

为了考察空间 $W^{m,p}(\Omega)$ 上的连续线性泛函, 我们把 $W^{m,p}(\Omega)$ 等距嵌入到一个乘积空间 $\Pi_{|\alpha| \leqslant m} L^p(\Omega)$ 中去, 使之成为该空间的一个线性闭子空间,

$$\begin{aligned} i: \quad & u \mapsto \{D^\alpha u, |\alpha| \leqslant m\}, \\ W^{m,p}(\Omega) & \to \prod_{|\alpha| \leqslant m} L^p(\Omega). \end{aligned}$$

由 Hahn-Banach 定理, 为了 $f \in (W^{m,p}(\Omega))^*$ 必须且仅需存在 $\{\psi_\alpha, |\alpha| \leqslant m\} \in \prod_{|\alpha| \leqslant m} L^{p'}(\Omega)$ 使得

$$\langle f, u \rangle = \int_\Omega \left(\sum_{|\alpha| \leqslant m} D^\alpha u(x) \psi_\alpha(x) \right) dx.$$

于是 $W^{m,p}(\Omega), p \geqslant 1$ 中的弱收敛 $u_j \rightharpoonup u$ 可以表达为

$$u_j \rightharpoonup u \text{ 于 } W^{m,p}(\Omega) \Longleftrightarrow \int_\Omega \sum_{|\alpha| \leqslant m} D^\alpha (u_j - u) \psi_\alpha dx \to 0, \quad \forall \{\psi_\alpha\} \in \prod_{|\alpha| \leqslant m} L^{p'}(\Omega).$$

定理 10.2 $W^{m,p}(\Omega)(1 < p < \infty)$ 是自反的 Banach 空间.

证明 定义嵌入映射 $i : W^{m,p}(\Omega) \to \prod_{|\alpha| \leqslant m} L^p(\Omega)$ 为

$$i : u \mapsto \{D^\alpha u, |\alpha| \leqslant m\}.$$

它是闭的, 从而 $i(W^{m,p}(\Omega))$ 是 $\prod_{|\alpha| \leqslant m} L^p(\Omega)$ 的闭线性子空间. 因为 $\prod_{|\alpha| \leqslant m} L^p(\Omega)$ 是自反的, 所以由 Pettis 定理, $W^{m,p}(\Omega)$ 是自反的. □

§10.4 光滑化算子

在第六讲, 在证明高维的 du Bois-Reymond 引理时, 我们引入过钟形函数

$$\varphi(x) = \begin{cases} c_n^{-1} \exp\left(\dfrac{-1}{1 - |x|^2} \right), & |x| < 1, \\ 0, & |x| \geqslant 1, \end{cases}$$

其中 $|x| = (x_1^2 + \cdots + x_n^2)^{1/2}$, 而 c_n 是一个常数使得

$$\int_{\mathbb{R}^n} \varphi(x) dx = 1.$$

$\forall \varepsilon > 0$, 再令

$$\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi\left(\frac{x}{\varepsilon}\right),$$

则 $\mathrm{supp}\varphi_\varepsilon \subset \overline{B_\varepsilon(\theta)}$.

我们利用它来光滑化任意函数. 设 $u \in L^1(\Omega)$, $\mathrm{supp}\,(u) \subset \Omega_\delta := \{x \in \Omega \,|\, d(x, \partial\Omega) \geqslant \delta > 0\}$, 我们把映射 $u \mapsto u_\varepsilon$ $(0 < \varepsilon < \delta)$ 称为光滑化算子, 其中

$$u_\varepsilon(x) = \int_\Omega \varphi_\varepsilon(x-y)u(y)dy.$$

光滑化算子有下列简单性质.

(1) $\mathrm{supp}\,(u_\varepsilon) \subset (\mathrm{supp}\,u)_\varepsilon := \{x \in \Omega \,|\, d(x, \mathrm{supp}\,u) \leqslant \varepsilon\}$.

(2) $u_\varepsilon \in C_0^\infty(\Omega)$,

$$\partial^\alpha u_\varepsilon(x) = \int_\Omega u(x)\partial^\alpha \varphi_\varepsilon(x-y)dy.$$

(3) 若 $u \in C_0^m(\Omega)$, 则 $\partial^\alpha(u_\varepsilon) = (\partial^\alpha u)_\varepsilon$, $\forall\, \alpha$, $|\alpha| \leqslant m$.

(4) 若 $u \in C_0(\Omega)$, 则当 $\varepsilon \to 0$ 时, $\|u - u_\varepsilon\|_C \to 0$.

(5) 若 $u \in L^p(\Omega)$, $1 \leqslant p < \infty$, 则当 $\varepsilon \to 0$ 时, $\|u - u_\varepsilon\|_{L^p} \to 0$.

(6) $C_0^\infty(\Omega)$ 在 $L^p(\Omega)$ 内稠密.

(7) 设 $u, \partial^\alpha u \in L^p(\Omega)$, $p \in [1, \infty]$, $\mathrm{supp}\,u \subset \mathrm{int}(\Omega)$, 则当 ε 足够小时有 $(D^\alpha u)_\varepsilon = \partial^\alpha(u)_\varepsilon$.

事实上,

$$\begin{aligned}
\text{左式} &= \int_\Omega D^\alpha u(y)\varphi_\varepsilon(x-y)dy \\
&= (-1)^{|\alpha|}\int_\Omega u(y)\partial_y^\alpha \varphi_\varepsilon(x-y)dy \\
&= \int_\Omega u(y)\partial_x^\alpha \varphi_\varepsilon(x-y)dy \\
&= \partial_x^\alpha \int_\Omega u(y)\varphi_\varepsilon(x-y)dy = \text{右式}.
\end{aligned}$$

(8)

$$W_0^{m,p}(\mathbb{R}^n) = W^{m,p}(\mathbb{R}^n).$$

§10.5 Sobolev 空间的重要性质与嵌入定理

Sobolev 空间是非常基本的函数空间, 在调和分析、偏微分方程、泛函分析和变分学中都有基本的重要性, 它的重要性质在许多教科书中都有专门的介绍. 在本

节中, 我们仅罗列其中主要的结论, 详细的证明请参阅有关的书籍和文献. 不过为了帮助读者理解这些结果的意义和证明的实质, 我们还是愿意在一些特殊的或者比较简单的情形, 证明这些结论.

延拓定理

Sobolev空间是函数 (类) 空间, 这些函数可以定义在任意区域 $\Omega \subset \mathbb{R}^n$ 上, 函数空间是 $W^{m,p}(\Omega)$; 也可以定义在全空间 $\mathbb{R}^n$, 函数空间是 $W^{m,p}(\mathbb{R}^n)$. 自然的一个问题是: 是否每个 $W^{m,p}(\Omega)$ 中的函数 u 都可以扩张为 $\mathbb{R}^n$ 上的函数 $\tilde{u}$, $\tilde{u}|_\Omega = u$, 使得 $\tilde{u} \in W^{m,p}(\mathbb{R}^n)$?

如果 $\Omega \subset \mathbb{R}^n$ 有充分光滑的边界, 那么结论是肯定的.

定理 10.3 (延拓定理) *设 Ω 是一个有界区域, 且 $\partial\Omega$ 是一致 C^m 的, 则 $\forall 0 \leqslant l \leqslant m$, $\forall 1 \leqslant p < \infty$, 存在 $T \in L(W^{l,p}(\Omega), W^{l,p}(\mathbb{R}^n))$, 使得*

$$Tu(x) = u(x), \quad \text{a.e. } x \in \Omega.$$

证明参看 [Ad] 定理 4.26. 这个定理属于 Lichtenstein, Hestenes, Seeley 和 Calderon.

但是对于 $W_0^{m,p}$ 型空间, 不论区域 Ω 如何, 延拓总是可能的, 即总存在 $T \in L(W_0^{l,p}(\Omega), W_0^{l,p}(\mathbb{R}^n))$, 使得

$$Tu(x) = \begin{cases} u(x), & \text{a.e. } x \in \Omega, \\ 0, & x \notin \Omega. \end{cases}$$

这是因为 $C_0^\infty(\Omega)$ 是 $W_0^{m,p}(\Omega)$ 的稠密子空间, 而 $C_0^\infty(\Omega)$ 是 $C_0^\infty(\mathbb{R}^n)$ 的子空间, 从而 $W_0^{l,p}(\Omega)$ 是 $W_0^{l,p}(\mathbb{R}^n)$ 的闭子空间.

逼近定理

我们曾把 $H^1(\Omega)$ 定义为 $C^1(\bar{\Omega})$ 在模 $\|u\| = \left(\int_\Omega (|\nabla u|^2 + |u|^2)dx\right)^{1/2}$ 下的完备化空间. 并且 $W^{1,2}(\Omega)$ 是包含 $C^1(\bar{\Omega})$ 的完备空间, 因此 $H^1(\Omega) \subset W^{1,2}(\Omega)$. 我们自然要问: 它们是否相同? 以下逼近定理给出肯定的答复.

定理 10.4 (Serrin-Meyers 逼近定理) *若 $p \in [1, \infty)$, 则集合 $S := C^\infty(\Omega) \cap W^{m,p}(\Omega)$ 在 $W^{m,p}(\Omega)$ 中稠密.*

证明 选取一列开区域 $\{\Omega_j\}$ 满足

$$\varnothing = \Omega_{-1} = \Omega_0 \subset \Omega_1 \subset \bar{\Omega}_1 \subset \Omega_2 \subset \bar{\Omega}_2 \subset \cdots \subset \Omega_i \subset \bar{\Omega}_i \subset \Omega_{i+1} \subset \cdots,$$

$$\bigcup_{i=1}^\infty \Omega_i = \Omega.$$

例如说, $\Omega_i = \{x \in \Omega \mid \|x\| \leqslant i, \mathrm{dist}(x, \partial\Omega) > 1/i\}$, $i = 1, 2, \cdots$. 对于 Ω 的一族开覆盖

$$O = \{U_k = \Omega_{k+1} \backslash \Omega_{k-1} \mid k = 1, 2, \cdots\}$$

有对应的一族单位分解 $\{\psi_i\}$, 即

(1) $\psi_i \in C_0^\infty(\Omega)$,

(2) $\mathrm{supp}\{\psi_i\} \subset \Omega_{i+1} \backslash \Omega_{i-1}$,

(3) $\sum \psi_i(x) \equiv 1, \forall x \in \Omega$.

如令 $\forall u \in W^{m,q}(\Omega)$, $\forall \varepsilon > 0$, 可以取 $\delta_i < \mathrm{dist}(\Omega_i, \partial\Omega_{i+1})$ 足够小使得

$$\|(\psi_i u)_{\delta_i} - \psi_i u\|_{m,q} < \frac{\varepsilon}{2^i}. \tag{10.2}$$

这是因为由 §10.4 性质 (7),(5) 和引理 10.1, $\forall \alpha$, $|\alpha| \leqslant m$, 我们有

$$\|\partial^\alpha (\psi_i u)_{\delta_i} - D^\alpha(\psi_i u)\|_q = \|(D^\alpha(\psi_i u))_{\delta_i} - D^\alpha(\psi_i u)\|_q \to 0.$$

对 (10.2) 求和,

$$\Big\|\sum_{i=1}^\infty (\psi_i u)_{\delta_i} - u\Big\|_{m,q} = \Big\|\sum_{i=1}^\infty (\psi_i u)_{\delta_i} - \sum_{i=1}^\infty (\psi_i u)\Big\|_{m,q} < \varepsilon.$$

因为 $\forall x \in \Omega$, $\sum_{i=1}^\infty (\psi_i u)_{\delta_i}(x)$ 是有限和, 所以 $v = \sum_{i=1}^\infty (\psi_i u)_{\delta_i} \in C^\infty(\Omega)$ 是无穷次可微的.

又 $\forall k \in \mathbb{N}$, 在每个 Ω_k 上,

$$u(x) = \sum_{i=1}^{k+2} (\psi_i u)(x), \quad v(x) = \sum_{i=1}^{k+2} (\psi_i u)_{\delta_i}(x),$$

有估计

$$\|u - v\|_{W^{m,p}(\Omega_k)} \leqslant \sum_{i=1}^{k+2} \|(\psi_i u)_{\delta_i} - \psi_i u\|_{m,p} < \varepsilon.$$

再令 $k \to \infty$, 即得 $\|u - v\|_{m,p} < \varepsilon$. 显然 $v \in S$, 定理证完. □

推论 10.1 如果 Ω 是一个具有一致 C^m 边界的有界区域, 那么 $\forall 1 \leqslant p < \infty$, $W^{m,p}(\Omega)$ 是可分的 Banach 空间.

证明 取定一个开立方体 $D \subset \mathbb{R}^n$ 使得 $\Omega \subset \bar{\Omega} \subset D$. $\forall u \in W^{m,q}(\Omega), \exists\, \tilde{u} \in W^{m,q}(\mathbb{R}^n)$, 使得

$$\|\tilde{u}\|_{W^{m,q}(\mathbb{R}^n)} \leqslant C\|u\|_{W^{m,q}(\Omega)}, \quad \tilde{u}|_\Omega = u.$$

对于任意的 $\varepsilon > 0$, 应用定理 10.4, $\exists v \in C^\infty(\mathbb{R}^n)$ 使得

$$\|\tilde{u} - v\|_{W^{m,q}(\Omega)} \leqslant \|\tilde{u} - v\|_{W^{m,q}(D)} \leqslant \frac{\varepsilon}{3}.$$

取 $\psi \in C_0^\infty(D)$ 满足 $\psi(x) = 1, \forall x \in \Omega$. 函数 $\psi v \in C_0^\infty(D)$. 再由 Weierstrass 多项式逼近定理存在 $\bar{D}$ 上的多项式 P 满足

$$\|\psi v - P\|_{W^{m,q}(D)} \leqslant \|\psi v - P\|_{C^m(\bar{D})} \leqslant \frac{\varepsilon}{3}.$$

有理系数的多项式可以 C^m 逼近任意多项式, 这就证明了 $W^{m,q}(\Omega)$ 存在可数稠密集. □

推论 10.2 $H^1(\Omega) = W^{1,2}(\Omega)$.

Poincaré 不等式

第九讲中的 Poincaré 不等式可以推广到 $W_0^{1,p}(\Omega), 1 \leqslant p < \infty$.

Poincaré 不等式 设 $\Omega \subset \mathbb{R}^n$ 是有界区域. $u \in W_0^{1,p}(\Omega), 1 \leqslant p < \infty$, 则 $\exists C = C(p, \Omega)$ 使得

$$\int_\Omega |u|^p dx \leqslant C \int_\Omega |\nabla u|^p dx.$$

推论 10.3 设 $\Omega \subset \mathbb{R}^n$ 是有界区域, 则

$$\|u\| = \left(\int_\Omega |\nabla u|^p\right)^{1/p}$$

是 $W_0^{1,p}(\Omega)$ $(1 \leqslant p < \infty)$ 空间的一个等价范数.

$W_0^{1,p}(\Omega)$ 是 $W^{1,p}(\Omega)$ 的线性闭子空间, 从而也是自反的 Banach 空间.

嵌入定理

定理 10.5 (Sobolev) 嵌入映射

$$W^{m,q}(\mathbb{R}^n) \hookrightarrow L^r(\mathbb{R}^n), \quad \frac{1}{r} = \frac{1}{q} - \frac{m}{n} \ (mq < n),$$

以及 $\forall j \in \mathbb{N}$,

$$W^{m+j,q}(\mathbb{R}^n) \hookrightarrow C^{j,\lambda}(\mathbb{R}^n), \quad 0 < \lambda \leqslant m - \frac{n}{q} \ (mq > n)$$

都是连续的.

证明参看 [Ad] 5.4–5.10.

联合上述嵌入定理与延拓定理, 我们有

定理 10.6 (Sobolev 嵌入定理) 设 Ω 是一个具有一致 C^m 边界的有界区域, $1 \leqslant q < \infty, m \geqslant 0$ 是整数, 则嵌入映射

$$W^{m,q}(\Omega) \hookrightarrow L^r(\Omega), \quad \frac{1}{r} \geqslant \frac{1}{q} - \frac{m}{n} \ (mq < n),$$

以及 $\forall j \in \mathbb{N}$,

$$W^{m+j,q}(\Omega) \hookrightarrow C^{j,\lambda}(\bar{\Omega}), \quad 0 < \lambda \leqslant m - \frac{n}{q} \ (mq > n)$$

都是连续的.

最常用的是 $m = 1$ 的情形, 记 $q^* = \dfrac{nq}{n-q}$, 则

$$W^{1,q}(\Omega) \hookrightarrow L^r(\Omega), \quad r \leqslant q^* \ (n > q),$$

以及

$$W^{1,q}(\Omega) \hookrightarrow C(\bar{\Omega}) \quad (q > n)$$

都是连续的.

注 10.1 当 $n = 1, \Omega = (a, b)$ 时, 嵌入定理的结论很容易从 Hölder 不等式和例 10.3 直接推出来. 因为此时广义导数与几乎处处导数是一致的, 所以

$$|u(x)-u(y)| = \left|\int_y^x u'(t)dt\right| \leqslant |x-y|^{1/q'} \left(\int_a^b |u'(t)|^q dt\right)^{1/q}, \quad \forall\, x, y \in (a, b). \ \square$$

紧嵌入

定理 10.7 (Rellich-Kondrachov 紧嵌入定理) 设 $\Omega \subset \mathbb{R}^n$ 是一个具有一致 C^m 边界的有界区域, $1 \leqslant p \leqslant \infty, m \geqslant 0$ 是整数, 则嵌入映射

$$W^{m,q}(\Omega) \hookrightarrow L^r(\Omega), \quad 1 \leqslant r < \frac{nq}{n-mq} \left(m < \frac{n}{q}\right),$$

以及

$$W^{m,q}(\Omega) \hookrightarrow C(\bar{\Omega}) \left(m > \frac{n}{q}\right)$$

都是紧的.

证明参看 [Ad] 定理 6.2.

我们对下列常用的特殊情形给予直接的证明.

定理 10.8 (Rellich) 设 $\Omega \subset \mathbb{R}^n$ 是一个有界区域, $1 \leqslant p < \infty$, 则 $W_0^{1,p}(\Omega)$ 中的有界闭球是 $L^p(\Omega)$ 中的列紧集.

证明 只需证明单位闭球是列紧集. 记 B 为 $W_0^{1,p}(\Omega)$ 的中心在 0 的单位闭球. 我们要证明 $\forall \varepsilon > 0$ 在 L^p 模下, B 存在有限 ε-网.

我们的想法是在 B 的任意一个 L^p 邻域内, 寻求连续函数空间上的一个一致有界、等度连续的集合.

1. 由定义, $C_0^\infty(\Omega)$ 在 $W_0^{1,p}(\Omega)$ 中稠密. 记 $S = C_0^\infty(\Omega) \cap B$. 又对任意 $\delta > 0$, 记 $S_\delta = \{v_\delta \,|v \in S\}$, 其中

$$v_\delta(x) = \int_\Omega v(y)\varphi_\delta(x-y)dy.$$

$\forall\, v \in S$,

$$\|v_\delta\|_{1,p} \leqslant \|v\|_{1,p} \leqslant 1,$$

由于

$$\begin{aligned}|v_\delta(x) - v(x)| &= \left|\int_{|y|\leqslant 1} \varphi(y)[v(x) - v(x-\delta y)]dy\right| \\ &\leqslant \int_{|y|\leqslant 1} \varphi(y) \int_0^{\delta|y|} \left|\frac{\partial}{\partial r} v(x - ry/|y|)\right| drdy,\end{aligned}$$

利用 Hölder 不等式,

$$\begin{aligned}&\|v_\delta - v\|_p^p \leqslant \int_\Omega \left(\int_{|y|\leqslant 1} \varphi(y)^{1/p'}\varphi(y)^{1/p} \int_0^{\delta|y|} \left|\frac{\partial}{\partial r} v(x - ry/|y|)\right| drdy\right)^p dx \\ &\leqslant \int_{|y|\leqslant 1} \varphi(y)\delta^p|y|^p \int_\Omega |\nabla v(x)|^p dxdy,\end{aligned}$$

即

$$\|v_\delta - v\|_p \leqslant \delta\|\nabla v\|_p.$$

因此 $\exists \delta_0 = \delta(\varepsilon)$, 使得

$$\|v - v_\delta\|_p \leqslant \frac{\varepsilon}{8},\ \ \forall v \in S, \forall \delta \leqslant \delta_0.$$

固定 $\delta = \delta_0$, 我们有

$$|v_\delta(x)| = \left|\int_\Omega v(y)\varphi_\delta(x-y)dy\right| \leqslant \frac{C}{\delta^n}\|v\|_p,$$

以及

$$|\nabla v_\delta(x)| = \left|\int_\Omega v(y)\nabla\varphi_\delta(x-y)dy\right| \leqslant \frac{C}{\delta^{n+1}}\|v\|_p.$$

可见 $S_{\delta_0} \subset C(\bar\Omega)$ 是一致有界、等度连续的. 根据 Arzelà-Ascoli 定理, 在 C 模下, S_{δ_0} 存在有限的 $\frac{\varepsilon}{4\mathrm{mes}(\Omega)}$-网: $\{w_1, w_2, \cdots, w_l\}$. 即, $\forall v \in S_{\delta_0}, \exists w_i \in S_{\delta_0}$ 使得

$$\|v_\delta - w_i\|_C < \frac{\varepsilon}{4\mathrm{mes}(\Omega)}.$$

于是

$$\|v_\delta - w_i\|_p < \frac{\varepsilon}{4}.$$

2. 然而 w_i 未必在 B 内, 但对每个 w_i 对应着一个 $v^i_{\delta_0}$, 其中 $v^i \in S = C_0^\infty(\Omega)\cap B$. 如果 $v^i_{\delta_0} \notin B$, 即它的支集超出了 $\bar\Omega$, 那么我们可以取 $\delta_i \in (0, \delta_0]$, 使得 $w_i' = v^i_{\delta_i}$ 的支集落到 Ω 内, 这时便有 $w_i' \in B$, 以及

$$\|w_i - w_i'\|_p \leqslant \|w_i - v^i\|_p + \|w_i' - v^i\|_p \leqslant \frac{\varepsilon}{4}.$$

如今 $\forall u \in B, \exists v \in S$ 使得 $\|u - v\|_{1,p} \leqslant \varepsilon/4$, 所以有

$$\|u - w_i'\|_p \leqslant \|u - v\|_p + \|v - v_{\delta_0}\|_p + \|v_{\delta_0} - w_i\|_p + \|w_i - w_i'\|_p < \varepsilon.$$

这个 $\{w_1', \cdots, w_l'\}$ 就是我们所要的有限 ε-网. □

§10.6 Euler-Lagrange 方程

我们要把前面在 C^1 空间上讨论过的变分理论推广到 Sobolev 空间. 给定一个 Lagrange 函数 $L \in C(\Omega \times \mathbb{R}^N \times \mathbb{R}^{nN}), L = L(x, u, p)$ 是可微的, 其中 $\Omega \subset \mathbb{R}^n$. 为了使泛函

$$I(u) = \int_\Omega L(x, u(x), \nabla u(x))dx$$

在某个 Sobolev 空间 $W^{1,q}(\Omega)$ 上有定义, 并且极值的必要条件 E-L 方程也有意义, 我们还需要对 L 添加如下假设:

(1) L, L_u, L_p 是连续的;

(2) $|L(x, u, p)| \leqslant C(1 + |u|^q + |p|^q)$;

(3) $|L_u(x, u, p)| + |L_p(x, u, p)| \leqslant C(1 + |u|^q + |p|^q)$.

事实上, 由假设 (2), $\forall\, u \in W^{1,q}(\Omega, \mathbb{R}^N)$,

$$I(u) = \int_\Omega |L(x, u(x), \nabla u(x))|dx \leqslant C\int_\Omega (1 + |u(x)|^q + |\nabla u(x)|^q)dx < \infty.$$

再由假设 (3),

$$\int_\Omega |L_u(x,u(x),\nabla u(x))|^q + |L_p(x,u(x),\nabla u(x))|^q dx < \infty,$$

即函数 $\lambda(x) := L_u(x,u(x),\nabla u(x))$ 与 $\mu(x) := L_p(x,u(x),\nabla u(x)) \in L^q$.

引理 10.2 在 (1), (2), (3) 的假设下, $\forall\, u_0 \in W^{1,q}(\Omega)$, $\forall \varphi \in C_0^1(\Omega)$,

$$\begin{aligned}\delta I(u_0,\varphi) &= \frac{d}{dt} I(u_0 + s\varphi)|_{s=0} \\ &= \int_\Omega [L_u(x,u(x),\nabla u(x))\varphi(x) + L_p(x,u(x),\nabla u(x))\nabla\varphi(x)]dx. \end{aligned} \tag{10.3}$$

证明 我们只需关注

$$s^{-1}[I(u+s\varphi) - I(u)] = \int_\Omega [\lambda_s(x)\varphi(x) + \mu_s(x)\nabla\varphi(x)]dx,$$

其中

$$\lambda_s(x) = \int_0^1 L_u(x,u(x)+\theta s\varphi(x), \nabla(u(x)+\theta s\varphi(x)))d\theta,$$
$$\mu_s(x) = \int_0^1 L_p(x,u(x)+\theta s\varphi(x), \nabla(u(x)+\theta s\varphi(x)))d\theta,$$

由假设 (3), $\lambda_s, \mu_s \in L^{q'}$, $q^{-1} + q'^{-1} = 1$, 并且当 $s \to 0$ 时, 几乎处处

$$\lambda_s(x) \to L_u(x,u(x),\nabla u(x)), \quad \mu_s(x) \to L_p(x,u(x),\nabla u(x)).$$

此外, $\forall \varphi \in W_0^{1,q}(\Omega,\mathbb{R}^N)(C_0^\infty(\Omega,\mathbb{R}^N))$, (10.3) 的积分号内有可积的控制函数

$$C(1+|u(x)|^{q-1} + |\nabla u(x)|^{q-1})(|\varphi(x)| + |\nabla\varphi(x)|).$$

应用 Lebesgue 控制定理, 即得结论. □

注 10.2 引理 10.2 的假设 (1) 中对三个变量 (x,u,p) 的连续性可以换成较弱的 Carathéodory 条件:

$$\begin{aligned}&\forall (u,p) \in \mathbb{R}^N \times \mathbb{R}^{nN},\ x \mapsto L(x,u,p) \text{ 是可测的}, \\ &\text{对于 a.e. } x \in \Omega,\ (u,p) \mapsto L(x,u,p) \text{ 是连续的}.\end{aligned} \tag{10.4}$$

即, 设 L, L_u, L_p 满足 Carathéodory 条件. □

注 10.3 注意到 Sobolev 嵌入定理: 当 $q < n$ 时, $W^{1,q}(\Omega) \hookrightarrow L^r(\Omega), r \leqslant q^* = \frac{nq}{n-q}$. 引理 10.2 的假设 (3) 中的增长幂次可以放松为:

$$(3')\ |L_u(x,u,p)| + |L_p(x,u,p)| \leqslant \begin{cases} C(1+|u|^r+|p|^q), & r \leqslant q^*,\ q < n, \\ C(1+|u|^r+|p|^q), & r \geqslant 1,\ \ q = n, \\ C(1+|p|^q), & q > n. \end{cases} \qquad \square$$

于是得到如下定理.

定理 10.9 设 Carathéodory 条件 (10.4) 与假设 (2) 和 (3′) 成立. 给定 $\rho \in W^{1,q}(\Omega, \mathbb{R}^N), u_0 \in M = \rho + W_0^{1,q}(\Omega, \mathbb{R}^N)$ 是 I 在 M 中的极小点, 则 u_0 满足如下形式的 E-L 方程:

$$\int_\Omega [L_u(x,u(x),\nabla u(x))\varphi(x) + L_p(x,u(x),\nabla u(x))\nabla\varphi(x)]dx = 0,$$

$$\forall \varphi \in C_0^\infty(\Omega, \mathbb{R}^N). \qquad (10.5)$$

在这个意义上, 我们把满足 (10.5) 的解称为 E-L 方程的广义解.

注 10.4 方向导数的概念在 Banach 空间的微分学中对应着 Gâteaux 导数.

设 X 是一个 Banach 空间. 设 $U \subset X$ 是一个开集, 给定 U 上的一个函数 $f \in C(U, \mathbb{R}^1)$. $\square$

定义 10.4 (Gâteaux 导数) 设 $x_0 \in U$, 我们称 f 在 x_0 是 Gâteaux 可微的, 如果 $\forall h \in X, \exists df(x_0,h) \in \mathbb{R}^1$ 使得

$$|f(x_0+th) - f(x_0) - tdf(x_0,h)| = o(t) \quad (t \to 0).$$

我们称 $df(x_0,h)$ 为 f 在 x_0 处的 Gâteaux 导数.

在变分学中, 对于空间 $W^{1,q}(\Omega, \mathbb{R}^N)$ 上的一个泛函 I 来说, 变分 $\delta I(u_0,\varphi)$ 与 Gâteaux 导数 $dI(u_0,\varphi)$ 的区别在于: 在变分中, $\varphi \in C_0^\infty(\Omega, \mathbb{R}^N)$ 或 $C_0^1(\Omega, \mathbb{R}^N)$, 在 Gâteaux 导数中, $\varphi \in W_0^{1,q}(\Omega, \mathbb{R}^N)$.

为了 $dI(u_0,\varphi)$ 中的积分有意义, 我们必须把 (3) 中的增长条件修改为

$$(3'')\ |L_u(x,u,p)| + |L_p(x,u,p)| \leqslant \begin{cases} C(1+|u|^{r-1}+|p|^{q-1}), & r < q^*,\ q < n \\ C(1+|u|^r+|p|^{q-1}), & r \geqslant 1,\ \ q = n \\ C(1+|p|^{q-1}), & q > n. \end{cases}$$

对于变分积分 I 的 Gâteaux 导数有如下定理.

定理 10.10 假设 (1), (2), (3″) 成立. 给定 $\rho \in W^{1,q}(\Omega,\mathbb{R}^N)$. 如果 $M = \rho + W_0^{1,q}(\Omega,\mathbb{R}^N)$ 上的泛函: $I(u) = \int_\Omega L(x,u(x),\nabla u(x))dx$ 在 $u_0 \in M$ 达到极小值, 那么 u_0 是 Gâteaux 可微的, 并且

$$dI(u_0,\varphi) = \int_\Omega [L_u(x,u(x),\nabla u(x))\varphi(x) + L_p(x,u(x),\nabla u(x))\nabla\varphi(x)]dx,$$
$$\forall \varphi \in W_0^{1,q}(\Omega,\mathbb{R}^N).$$

比较 (3′) 与 (3″), 显然 (3″) 更强, 所以在变分学中我们一般不用 Gâteaux 导数, 而直接用变分导数.

注 10.5 当 $n = 1, J = (a,b)$ 时, 利用 (10.5) 还能导出积分形式的 E-L 方程:

$$\int_a^t L_u(t,u^*(t),\dot{u}^*(t))dt - L_p(t,u^*(t),\dot{u}^*(t)) = c, \quad \text{a.e.}. \tag{10.6}$$

在第二讲的注 2.2 中我们已经在 $L^\infty(J)$ 上建立了 du Bois-Reymond 引理, 我们现在还要把这个引理推广到 Lebesgue 空间 $L^1(J)$. □

引理 10.3 设 $f \in L^1(J)$ 满足

$$\int_J f(t)\varphi'(t)dt = 0, \quad \forall \varphi \in C_0^\infty(J),$$

则 $f(t) = c$, a.e. $t \in J$.

证明 固定 f 的任意两个 Lebesgue 点 $a < t_0 < t_1 < b$. 令 $c = f(t_0)$, 取 $\varepsilon > 0$ 使得 $(t_0-\varepsilon, t_1+\varepsilon) \subset J$. 构造一个逐段线性函数 ψ:

$$\psi(t) = \begin{cases} 0, & t \notin [t_0-\varepsilon, t_1+\varepsilon], \\ 1, & t \in [t_0, t_1]. \end{cases}$$

其余部分用直线段连接. 显然存在 $\{\varphi_n\} \subset C_0^\infty(J)$, 使得 $\varphi_n \to \psi$, $W^{1,1}(J)$. 那么

$$\int_J f(t)\varphi_n'(t)dt \to \int_J f(t)\psi'(t)dt,$$

从而

$$\varepsilon^{-1}\int_{t_0-\varepsilon}^{t_0} f(t)dt - \varepsilon^{-1}\int_{t_1}^{t_1+\varepsilon} f(t)dt = 0.$$

当 $s \to 0$ 时, 取极限得到 $f(t_1) = f(t_0) = c$. 因为 t_1 可以取成任意 Lebesgue 点, 我们得到, 对于一切 Lebesgue 点 $t \in J, f(t) = c$. 引理证毕. □

于是在假设 (1), (2), (3′) 下, I 的极小点 $u^* \in u_0 + W_0^{1,q}(\Omega,\mathbb{R}^N)$ 满足积分形式的 E-L 方程 (10.6).

习 题

1. 验证 10.4 一节中的简单性质 (1)–(6), (8).

2. 设 $u \in W^{1,q}(\mathbb{R}^1)$, 证明:

$$u \in \begin{cases} C^\lambda(\mathbb{R}^1), \ \lambda = \dfrac{q-1}{q}, & 1 \leqslant q < \infty, \\ \mathrm{Lip}(\mathbb{R}^1), & q = \infty. \end{cases}$$

第十一讲　弱下半连续性

在第九讲中, 我们指出过泛函的序列弱下半连续性在直接方法中起重要的作用. 在这一讲, 我们专门来讨论怎样判定泛函的序列弱下半连续性.

§11.1　凸集与凸函数

我们从集合的拓扑性质入手来讨论函数的下半连续性. 从定义可知, 在拓扑空间 X 上, 函数 $f : X \to \mathbb{R}^1$ 是下半连续的当且仅当它的下方水平集 $f_t = \{x \in X \mid f(x) \leqslant t, \forall t \in \mathbb{R}^1\}$ 是闭的. 在线性赋范空间 X 上, 函数 f 是序列弱下半连续的当且仅当集合 f_t $(\forall t \in \mathbb{R}^1)$ 是序列弱闭的.

一般说来, 集合的闭与弱闭是不同的概念, 弱闭集必是闭集; 但反过来未必成立.

在 Banach 空间中凸集有如下重要的性质: 设 $C \subset X$ 是 Banach 空间 X 的一个凸子集, 则

$$\text{闭} \Longleftrightarrow \text{序列弱闭}.$$

这是由于 Banach 空间有一个 Mazur 定理: 若 $x_n \rightharpoonup x$, 则 $\frac{1}{n}\sum_{i=1}^{n} x_i \to x$.

我们可以应用这个结果来判定泛函的序列弱下半连续性. 回忆凸函数的定义: 在线性空间 E 的一个凸子集 C 上给定的一个函数 f, 如果

$$f(\lambda x + (1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda) f(y), \quad \forall x, y \in C,\ \forall \lambda \in [0,1],$$

那么我们说它是凸的. 不难看出

$$f \text{ 是凸函数 } \Longrightarrow f_t = \{x \in X \mid f(x) \leqslant t\} \text{ 是凸集}, \ \forall t \in \mathbb{R}^1.$$

把这两个结论结合起来, 我们有如下定理.

定理 11.1 *设 X 是 Banach 空间, 若 $f: X \to \mathbb{R}^1$ 是凸函数, 则*

$$f \text{ 序列下半连续 } \Longleftrightarrow \text{ 序列弱下半连续}.$$

证明

$$f \text{ 下半连续 } \Longleftrightarrow \forall t \in \mathbb{R}^1, \ f_t \text{ 是闭集}.$$
$$f \text{ 序列弱下半连续 } \Longleftrightarrow \forall t \in \mathbb{R}^1, \ f_t \text{ 是序列弱闭集}.$$

因为 $\forall t \in \mathbb{R}^1$, f_t 是凸集, 所以二者等价. □

应用定理 11.1, 我们来判定一些泛函的序列弱下半连续性.

今后在不会发生误解时, 在记号上, 我们不再区分梯度 ∇ 与广义导数意义下的梯度 D, 一概都记作 ∇.

例 11.1 设 $\Omega \subset \mathbb{R}^n$ 是有界区域, $f \in L^2(\Omega)$. 在 Sobolev 空间 $H_0^1(\Omega)$ 上, 泛函

$$I(u) = \int_\Omega \left(\frac{1}{2}|\nabla u(x)|^2 + f(x)u(x)\right) dx$$

是序列弱下半连续的.

证明 因为 $u \to \displaystyle\int_\Omega f(x)u(x)dx$ 是线性的, 而且由 Poincaré 不等式

$$\int_\Omega f(x)u(x)dx \leqslant \|f\|_{L^2}\|u\|_{L^2} \leqslant C(f)\|u\|_{H_0^1},$$

它又是连续的, 此外它还是凸的, 所以这个泛函是序列弱下半连续的.

又因为 $u \to \int_\Omega |\nabla u(x)|^2 dx$ 是连续凸的, 所以也是序列弱下半连续的. □

例 11.2 设 $\Omega \subset \mathbb{R}^n$ 是有界区域, 在 Sobolev 空间 $W_0^{1,q}(\Omega)$ $(1 \leqslant q < \infty, 1 \leqslant r \leqslant q^* = \frac{nq}{n-q})$ 上, 泛函

$$I(u) = \int_\Omega (|\nabla u(x)|^q + |u(x)|^r) dx$$

是序列弱下半连续的. (同理, 留给读者自己验证.)

若换成

$$I(u) = \int_\Omega (|\nabla u(x)|^q + c(x)|u(x)|^r)dx,$$

其中 $c \in L^\infty(\Omega), c \geqslant 0$, 则结论不变, 证明也不变. □

要是去掉 c 的非负性限制, 一般来说, 因为 I 不再是凸的, 所以未必还有序列弱下半连续性. 然而如果再添加条件 $r < q^* = \frac{nq}{n-q}$, 那么 I 仍是序列弱下半连续的.

事实上, 若 $u_j \rightharpoonup u$, 于 $W_0^{1,q}(\Omega)$, 则

$$\liminf \int_\Omega |\nabla u_j(x)|^q dx \geqslant \int_\Omega |\nabla u(x)|^q dx.$$

我们还有

$$\liminf \int_\Omega c(x)|u_n(x)|^r dx \geqslant \int_\Omega c(x)|u(x)|^r dx.$$

因若不然, 则必存在 $\varepsilon > 0$, 以及子列 $\{u_{j'}\}$, 使得

$$\int_\Omega c(x)|u_{j'}(x)|^r dx < \int_\Omega c(x)|u(x)|^r dx - \varepsilon.$$

应用 Rellich-Kondrachov 定理有子列 $\{u_{n_j}\} \subset \{u_{j'}\}$, $u_{n_j} \to u$, 于 $L^r(\Omega)$, 所以有

$$\int_\Omega c(x)|u_{n_j}(x)|^r dx \to \int_\Omega c(x)|u(x)|^r dx.$$

这是一个矛盾! □

§11.2 凸性与弱下半连续性

为了变分积分 $I(u) = \int_\Omega L(x, u(x), \nabla u(x))dx$ 对于 u 是凸的, 需要假设 Lagrange 函数 $L(x, u, p)$ 对于 (u, p) 是凸的. 然而从例 11.2 中的讨论可知, 这个要求太强了. 利用紧嵌入定理, 对 u 的凸性要求可以被添加一定的增长性条件所取代. 我们要把这个想法应用到一般的变分问题中去.

定理 11.2 (Tonelli-Morrey) 设 $L : \bar{\Omega} \times \mathbb{R}^N \times \mathbb{R}^{nN} \to \mathbb{R}^1$, 满足

(1) $L \in C^1(\bar{\Omega} \times \mathbb{R}^N \times \mathbb{R}^{nN})$,

(2) $L \geqslant 0$,

(3) $\forall (x, u) \in \Omega \times \mathbb{R}^N, p \mapsto L(x, u, p)$ 是凸的,

则 $I(u)=\int_J L(x,u(x),\nabla u(x))dx$ 在 $W^{1,q}(\Omega,\mathbb{R}^N)$ $(1\leqslant q<\infty)$ 上是序列弱下半连续的.

现在我们要用到凸函数的另一个重要特征.

引理 11.1 设 X 是一个 Banach 空间, $h\in X$, 又设 $f:X\to\mathbb{R}^1$ 是一个凸函数. 如果 f 在 x_0 对于 h 有方向导数 $df(x_0,h)$, 那么

$$f(x_0)+df(x_0,h)\leqslant f(x_0+h).$$

这是单变量凸函数熟知结论的直接应用.

定理 11.2 的证明 设 $u_j\rightharpoonup u$ $(W^{1,q})$, 我们要证明

$$\liminf\int_\Omega L(x,u_j(x),\nabla u_j(x))dx\geqslant\int_\Omega L(x,u(x),\nabla u(x))dx.$$

1. 由 Rellich-Kondrachov 定理, 可以找到一个子列, 仍记作 $\{u_j\}$, 使得

$$u_j\rightharpoonup u,\quad L^q(\Omega,\mathbb{R}^N).$$

再由 Riesz 定理, 还有子列仍记作 $\{u_j\}$ 使得

$$u_j(x)\to u(x),\quad \text{a.e. } x\in\Omega.$$

$\forall\varepsilon>0$, 我们知道存在紧集 $K\subset\Omega$ 使得 $\operatorname{mes}(\Omega\backslash K)<\varepsilon$, 且
(1) 在 K 上一致地有 $u_j\rightrightarrows u$ (Egorov 定理),
(2) u 及 ∇u 在 K 上是连续的 (Luzin 定理),
(3) 若 $\int_\Omega L(x,u(x),\nabla u(x))dx<+\infty$, 则

$$\int_K L(x,u(x),\nabla u(x))dx\geqslant\int_\Omega L(x,u(x),\nabla u(x))dx-\varepsilon$$

(Lebesgue 积分的绝对连续性定理).

若 $\int_\Omega L(x,u(x),\nabla u(x))dx=+\infty$, 则取 $\int_K L(x,u(x),\nabla u(x))dx>\frac{1}{\varepsilon}$.

2. 因为 L 对 p 是凸的, 所以应用引理 11.1 得到

$$
\begin{aligned}
I(u_j) &\geqslant \int_K L(x,u_j(x),\nabla u_j(x))dx \\
&\geqslant \int_K L_p(x,u_j(x),\nabla u(x))(\nabla u_j(x)-\nabla u(x))dx + \int_K L(x,u_j(x),\nabla u(x))dx \\
&= \int_K L(x,u_j(x),\nabla u(x))dx + \int_K L_p(x,u(x),\nabla u(x))(\nabla u_j(x)-\nabla u(x))dx \\
&\quad + \int_K (L_p(x,u_j(x),\nabla u(x)) - L_p(x,u(x),\nabla u(x)))(\nabla u_j(x)-\nabla u(x))dx \\
&= \mathrm{I}+\mathrm{II}+\mathrm{III}.
\end{aligned}
$$

3. 由 (1), 在 K 上, $u_j \rightrightarrows u$, L 又是连续的, 所以

$$
\mathrm{I} = \int_K L(x,u_j(x),\nabla u(x))dx \to \int_K L(x,u(x),\nabla u(x))dx.
$$

由 (2), $L_p(x,u(x),\nabla u(x)) \in L^\infty(K)$, 取 χ_K 为 K 的特征函数, 于是

$$
\chi_K L_p(x,u(x),\nabla u(x)) \in L^\infty(\Omega) \subset L^{q'}(\Omega).
$$

利用 $u_j \rightharpoonup u\,(W^{1,q}(\Omega))$, 推出

$$
\nabla u_j \rightharpoonup \nabla u \ (L^q(\Omega,\mathbb{R}^{nN})),
$$

可见第二项

$$
\mathrm{II} = \lim_{j\to\infty} \int_K L_p(x,u(x),\nabla u(x))(\nabla u_j(x)-\nabla u(x))dx = 0.
$$

最后利用弱收敛列是有界列,

$$
\|\nabla u_j - \nabla u\|_1 \leqslant C_1(\|\nabla u_j - \nabla u\|_q) \leqslant C_1(\|\nabla u_j\|_q + \|\nabla u\|_q) \leqslant C_2,
$$

以及在 K 上一致地

$$
L_p(x,u_j(x),\nabla u(x)) \rightrightarrows L_p(x,u(x),\nabla u(x)),
$$

推得

$$
\mathrm{III} = \lim_{j\to\infty} \int_K (L_p(x,u_j(x),\nabla u(x)) - L(x,u(x),\nabla u(x)))(\nabla u_j(x)-\nabla u(x))dx = 0.
$$

总结如下:

$$
\varliminf_{j\to\infty} I(u_j) \geqslant \int_K L(x,u(x),\nabla u(x))dx \geqslant I(u) - \varepsilon.
$$

由 $\varepsilon > 0$ 的任意性即得

$$
\varliminf_{j\to\infty} I(u_j) \geqslant I(u).
$$

注 11.1 值得注意的是在定理中没有上方增长性限制，这是因为我们只涉及下半连续性，没有上方的增长性条件无非可能出现 $u \in W^{1,q}(\Omega)$ 使得 $I(u) = +\infty$，但这不影响下半连续性的讨论. □

利用 Carathéodory 函数的概念，定理 11.2 中的假设 (1) 可以换成

(1′) 对 a.e. (x, u)，$p \to L(x, u, p)$ 是可微的，L 连同 L_p 都是 Carathéodory 函数.

§11.3 一个存在性定理

定理 11.3 (极值的存在性) 设 $\Omega \subset \mathbb{R}^n$ 是一个有界可测集. 设 $u_0 \in W^{1,q}(\Omega, \mathbb{R}^N)$，$1 < q < \infty$，又设

(1) L, L_p 都是 Carathéodory 函数,

(2′) $\exists a \in C(\bar{\Omega}), b > 0$，使得 $L(x, u, p) \geqslant -a(x) + b|p|^q$，$\forall\,(x, u, p) \in \Omega \times \mathbb{R}^N \times \mathbb{R}^{nN}$,

(3) $\forall (x, u) \in \Omega \times \mathbb{R}^N, p \mapsto L(x, u, p)$ 是凸的,

则泛函

$$I(u) = \int_\Omega L(x, u(x), \nabla u(x)) dx$$

在 $u_0 + W_0^{1,q}(\Omega, \mathbb{R}^N)$ 上达到极小值.

证明 我们在自反的 Banach 空间 $W_0^{1,q}(\Omega, \mathbb{R}^N)$ 上考察泛函 $v \mapsto I(u_0 + v)$.

1. 条件 (2′) 与条件 (2) 虽然不尽相同，但可以用 $I + \int_\Omega a(x)dx$ 代替 I. 再应用 Tonelli-Morrey 定理于空间 $W_0^{1,q}(\Omega, \mathbb{R}^N)$，$I$ 是序列弱下半连续的.

2. I 是强制的，即

$$\text{当 } \|u\|_{1,q} \to \infty \text{ 时, } I(u) \to +\infty.$$

事实上，$\forall v \in W_0^{1,q}(\Omega, \mathbb{R}^N)$，由 Poincaré 不等式，存在 $\alpha, \beta > 0$ 使得

$$I(u_0 + v) \geqslant -\int_\Omega a(x)dx + b\int_\Omega |\nabla(u_0 + v)|^q dx \geqslant \alpha\|v\|_{1,q}^q - \beta.$$ □

作为一个例子，我们断言 Poisson 方程有广义解. 给定一个有界区域 $\Omega \subset \mathbb{R}^n$，给定 $f \in L^2(\Omega)$，存在 $u \in H_0^1(\Omega)$ 满足:

$$-\Delta u = f.$$

事实上，由定理 11.3，泛函

$$I(u) = \int_\Omega \left(\frac{1}{2}|\nabla u|^2 - fu\right) dx$$

有极小值.

§11.4 拟凸性*

虽然我们已经把凸性要求降低为 $\forall (x,u) \in \Omega \times \mathbb{R}^N, p \mapsto L(x,u,p)$ 是凸的, 但是在应用问题中这个要求有时还是太强了.

例如, Jacobi 行列式 $\det\left(\frac{\partial u^i}{\partial x_j}\right)$ 有明确的几何意义与力学意义. 当 $\Omega \subset \mathbb{R}^n$, $u = (u^1, u^2, \cdots, u^n), \Omega \to \mathbb{R}^n$ 时, 含有 u 的 Jacobi 行列式的 Lagrange 函数在几何与力学中会经常出现, 例如, $f \in C(\mathbb{R}^1, \mathbb{R}^1)$ 是一个凸函数,

$$\begin{aligned} L : \mathbb{R}^{n^2} &\to \mathbb{R}^1, \\ A \quad &\to L(A) = f(\det(A)), \end{aligned}$$

其中 $A \in \mathbb{R}^{n^2}$, 但 L 对 p 不是凸的.

我们自然要问, 前面所说的 L 对 p 的凸性假设是不是必须的? 我们先对 L 只依赖于 p, 而且 $n = 1$ 的简单情形来讨论, 得到启发后, 再考虑一般的情形.

设 $I(u) = \int_J L(u'(t))dt$ 在 $W^{1,\infty}(J)$ 上序列 * 弱下半连续, 即

$$u_j \rightharpoonup^* u(W^{1,\infty}(J, \mathbb{R}^N)) \Longrightarrow \underline{\lim} I(u_j) \geqslant I(u),$$

我们来寻求 L 应满足的条件.

不妨设 $J = (0,1)$, 特别地, $\forall \eta \in \mathbb{R}^N, \forall \lambda \in (0,1)$, 取函数

$$\varphi(t) = \begin{cases} t(1-\lambda)\eta, & t \in [0, \lambda), \\ (1-t)\lambda\eta, & t \in [\lambda, 1), \end{cases}$$

则 $\varphi \in W_0^{1,\infty}(J)$. 将其周期延拓到全直线, 令

$$\varphi_m(t) = \frac{1}{m}\varphi(mt).$$

现在对于任意的 $u_0, p_0 \in \mathbb{R}^N$, 作函数

$$u(t) = u_0 + p_0 t,$$

以及序列

$$u_m(t) = u(t) + \varphi_m(t), \quad m = 1, 2, \cdots.$$

我们有

$$|u_m(t) - u_m(s)| = |\varphi_m(t) - \varphi_m(s)| \leqslant \sup_{t \leqslant \theta \leqslant s} |\varphi'(m\theta)| \, |t - s|,$$

即

$$|u_m|_{1,\infty} \leqslant \|\varphi'\|_{L^\infty(J)},$$

而且

$$u_m \to u\ (\text{于 } L^\infty(J)).$$

再利用 $\dot{\varphi}_m(t) = \varphi'(mt)$, $\varphi \in W_0^{1,\infty}(J, \mathbb{R}^N)$, 以及 $\bar{\varphi}' = 0$. 按第九讲例 9.3, 可见

$$\varphi'_m \rightharpoonup^* 0\ (\text{于 } L^\infty(J)).$$

如今 $\forall \xi \in (W^{1,1}(J, \mathbb{R}^N))^*$, 我们有 $\xi = (\xi_0, \xi_1) \in L^1(J, \mathbb{R}^{2N})$, 又 $Du_m \rightharpoonup^* Du$ 于 $L^\infty(J)$, 便得到

$$\langle \xi, u_m - u \rangle = \int_J [\xi_0(u_m - u) + \xi_1 D(u_m - u)]dt \to 0.$$

即 $u_m \rightharpoonup^* u$, $W^{1,\infty}(J, \mathbb{R}^N)$.

注意到

$$I(u) = \int_J L(p_0)dt = L(p_0),$$

以及 φ 的周期性,

$$\begin{aligned} I(u_m) &= \int_J L(p_0 + \varphi'_m(t))dt \\ &= \frac{1}{m}\int_{mJ} L(p_0 + \varphi'(t))dt \\ &= \int_J L(p_0 + \varphi'(t))dt. \end{aligned}$$

假设 I 是序列弱下半连续的, $\liminf I(u_m) \geqslant I(u)$, 便推出

$$\int_J L(p_0 + \varphi'(t))dt \geqslant L(p_0).$$

有了前面的启发, 我们来考虑 $n > 1$ 的情形.

定理 11.4 设 $\Omega \subset \mathbb{R}^n$ 是一个区域, $L \in C(\Omega, \mathbb{R}^{nN})$. 如果

$$I(u) = \int_\Omega L(\nabla u(x))dx$$

在 $W^{1,\infty}(\Omega, \mathbb{R}^N)$ 上是 *弱序列下半连续的, 即

$$u_j \rightharpoonup^* u\ \ (W^{1,\infty}(\Omega, \mathbb{R}^N)) \Longrightarrow \underline{\lim} I(u_j) \geqslant I(u),$$

那么对于任意立方砖 $D \subset \bar{D} \subset \Omega$, $\forall A_0 \in \mathbb{R}^{nN}$($n \times N$ 矩阵),

$$\int_D L(A_0 + \nabla\varphi(x))dx \geqslant L(A_0)\mathrm{mes}(D),\ \forall \varphi \in W_0^{1,\infty}(\Omega, \mathbb{R}^N).$$

证明 1. 我们先把 φ_m 的构造推广到高维, 不妨设 $D = [0,1]^n$ 是单位立方砖. 对于任意 $k \in \mathbb{N}$, 把 D 作 2^k 等分, $D = \bigcup_{l=1}^{2^{kn}} D_l^k$, 其中 D_l^k 是边长为 2^{-k}, 中心在 $c_l^k = 2^{-k}\left(y_1^l + \frac{1}{2}, y_2^l + \frac{1}{2}, \cdots, y_n^l + \frac{1}{2}\right)$ 的子立方砖, 而 $(y_1^l, y_2^l, \cdots, y_n^l), l = 1, 2, \cdots, 2^{kn}$ 遍历格点 $(0, 1, 2, \cdots, 2^k - 1)^n$. 如今对于任意 $v \in W_0^{1,\infty}(D, \mathbb{R}^N)$, 将其周期扩张到 $\mathbb{R}^n$. 令

$$w_k(x) = \frac{1}{2^k} v(2^k(x - c_l^k)), \quad x \in D_l^k, \forall\, l = 1, 2, \cdots, 2^{kn}.$$

便有

$$\nabla w_k(x) = \nabla v(2^k(x - c_l^k)), \quad x \in D_l^k, \forall\, l = 1, 2, \cdots, 2^{kn},$$

以及

$$\begin{cases} w_k \to 0, & 于\ L^\infty(D, \mathbb{R}^N), \\ \nabla w_k \rightharpoonup^* 0, & 于\ L^\infty(D, \mathbb{R}^N). \end{cases}$$

2. 定义

$$u_k(x) = A_0 x + w_k(x),\ \ k = 1, 2, \cdots,$$

其中 w_k 定义如上, 并且令其在 D 外为 0. 可见

$$u_k \rightharpoonup^* u = A_0 x.$$

于是

$$\mathrm{mes}(\Omega)L(A_0) = I(u) \leqslant \liminf I(u_k),$$

并且

$$\begin{aligned} I(u_k) &= \int_D L(A_0 + \nabla w_k(x))dx + \int_{\Omega\setminus D} L(A_0)dx \\ &= \sum_{l=1}^{2^{kn}} \int_{D_l^k} L(A_0 + \nabla v(2^k(x - c_l^k)))dx + L(A_0)\mathrm{mes}(\Omega\setminus D) \\ &= \int_D L(A_0 + \nabla v(x))dx + L(A_0)\mathrm{mes}(\Omega\setminus D)). \end{aligned}$$

这就是我们要证明的. □

为此 Morrey 引入函数的拟凸性 (quasi-convexity) 概念.

定义 11.1 函数 L 称为是拟凸的, 如果 $\forall A \in \mathbb{R}^{nN}$($n\times N$ 矩阵), $\forall$ 立方砖 $D\subset\mathbb{R}^n, \forall v\in W_0^{1,\infty}(D,\mathbb{R}^N)$, 有

$$\mathrm{mes}(D)L(A)\leqslant\int_D L(A+\nabla v(x))dx.$$

拟凸性的重要意义在于它能保证泛函的序列弱下半连续性. 下列定理的证明很长, 有兴趣的读者请参看 [Da], pp.156–167.

定理 11.5 (Morrey-Acerbi-Fusco) 当 $1\leqslant p<\infty$ 时, 若

$$I(u)=\int_\Omega L(\nabla u)dx$$

在 $W^{1,p}(\Omega,\mathbb{R}^N)$ 是序列弱下半连续的 (或当 $p=\infty$ 时, 是序列 *弱下半连续的), 则 L 是拟凸的. 反过来, 添加如下增长性条件:

$$\begin{cases} |L(A)|\leqslant\alpha(1+|A|), & p=1,\\ -\alpha(1+|A|^q)\leqslant L(A)\leqslant\alpha(1+|A|^p), & 1\leqslant q<p<\infty,\\ |L(A)|\leqslant\eta(|A|), & p=\infty,\end{cases}$$

其中 $\alpha>0$ 是常数, 而 η 是一个不减连续函数. 若 L 是拟凸的, 则当 $1\leqslant p<\infty$ 时, I 在 $W^{1,p}(\Omega,\mathbb{R}^N)$ 是序列弱下半连续的 (或当 $p=\infty$ 时, 是序列 *弱下半连续的).

哪些函数是拟凸的? 设 $f:\mathbb{R}^1\to\mathbb{R}^1$ 是凸的, 以及 $u\in L^1(\Omega)$, 则由 Jessen 不等式

$$f\left(\frac{1}{\mathrm{mes}(\Omega)}\int_\Omega u(x)\right)dx\leqslant\frac{1}{\mathrm{mes}(\Omega)}\int_\Omega f(u(x))dx,$$

可见若 L 是凸函数, 则 L 必是拟凸函数. 事实上, 设 $p\to L(p)$ 是凸函数, 则有

$$\begin{aligned} L(p)&=L\left(\mathrm{mes}(D)^{-1}\int_D(p+\nabla\varphi(x))dx\right)\\ &\leqslant\mathrm{mes}(D)^{-1}\int_D L(p+\nabla\varphi(x))dx,\quad\forall\varphi\in W_0^{1,\infty}(\Omega,\mathbb{R}^N).\end{aligned}$$

值得注意的是当 $n=1$ 或 $N=1$ 时, 拟凸 = 凸.

我们先就 $n=1$ 的情形证明拟凸 $\Rightarrow$ 凸.

设 $\xi,\eta\in\mathbb{R}^N$, $\forall\lambda\in[0,1]$. 令

$$\xi_1=\xi+(1-\lambda)\eta,$$
$$\xi_2=\xi-\lambda\eta,$$

则

$$\xi = \lambda\xi_1 + (1-\lambda)\xi_2,$$
$$\eta = \xi_1 - \xi_2.$$

定义函数

$$\varphi(t) = \begin{cases} t(1-\lambda)\eta, & t \in [0,\lambda), \\ (1-t)\lambda\eta, & t \in [\lambda, 1), \end{cases}$$

代入拟凸条件得到

$$\begin{aligned} L(\xi) &\leqslant \int_0^1 L(\xi + \varphi'(t))dt \\ &= \int_0^\lambda L(\xi_1)dt + \int_\lambda^1 L(\xi_2)dt \\ &= \lambda L(\xi_1) + (1-\lambda)L(\xi_2). \end{aligned}$$

这表明 $p \mapsto L(p)$ 是凸的.

我们再对 $N=1, n>1$ 的情形来证明拟凸 $\Rightarrow$ 凸, 即对 $\forall \xi_1, \xi_2 \in \mathbb{R}^n$, 证明

$$L(\lambda\xi_1 + (1-\lambda)\xi_2) \leqslant \lambda L(\xi_1) + (1-\lambda)L(\xi_2).$$

选取一个立方体 $D \subset \Omega$, 不妨设 $D = [0,1]^n$. 沿用上面的函数 φ, $\forall x = (x_1, \cdots, x_n) \in D$, 定义

$$u_k(x) = \eta k^{-1}\varphi(kx_1),$$

于是

$$\nabla u_k(x) = \begin{cases} (1-\lambda)\eta, & \{kx_1\} \in [0,\lambda), \\ -\lambda\eta, & \{kx_1\} \in [\lambda, 1), \end{cases}$$

其中 $\{y\}$ 表示 $y \in \mathbb{R}^1$ 的分数部分, $\eta = \xi_1 - \xi_2$. 再令

$$v_k(x) = \eta \min\{k^{-1}\varphi(kx_1), \operatorname{dist}(x, \partial D)\},$$

其中 $\operatorname{dist}(x,\partial D) = \inf\{\sup_{1\leqslant i\leqslant n}\|x_i - y_i\| \,|\, y = (y_1,\cdots,y_n) \in \partial D\}$. 可见 $v_k|_{\partial D} = 0$, 以及存在常数 $K > 0$ 使得 $|u_k(x) - v_k(x)| \leqslant K\|x - y\|$. 从而 $v_k \in W_0^{1,p}(D)$, $1 < p < \infty$. 此外,

$$\operatorname{mes}\{x \in D \,|\, \nabla u_k(x) \neq \nabla v_k(x)\} \to 0, \quad k \to \infty.$$

如今取 $D_1 = \{x \in D \,|\, \nabla u_k(x) = (1-\lambda)\eta\}$, $D_2 = \{x \in D \,|\, \nabla u_k(x) = -\lambda\eta\}$, 则 $D = D_1 \cup D_2$.

由于 L 是拟凸的, 所以

$$\text{mes}(D)L(\lambda\xi_1 + (1-\lambda)\xi_2) \leqslant \int_D L(\lambda\xi_1 + (1-\lambda)\xi_2 + \nabla v_k(x))dx.$$

令 $\xi = \lambda\xi_1 + (1-\lambda)\xi_2$, 则由积分的绝对连续性, 我们有

$$\begin{aligned}\lim \int_D L(\xi + \nabla v_k(x))dx &= \lim \int_D L(\xi + \nabla u_k(x))dx \\ &= \int_{D_1} L(\xi_1)dx + \int_{D_2} L(\xi_2)dx \\ &= (\lambda L(\xi_1) + (1-\lambda)L(\xi_2))\text{mes}(D).\end{aligned}$$

即得结论.

然而存在不是凸的拟凸函数. 例如, 矩阵的行列式

$$A \mapsto \det(A).$$

我们仅就 $N = n = 2$ 验证这个结论. 给定矩阵 $A = (a_{ij})$. 设 $u = (u_1, u_2)$, $P = (p^i_j)$, $L(P) = \det(P)$. 因为

$$\det(\nabla\varphi) = \partial_1(\varphi_1\partial_2\varphi_2) - \partial_2(\varphi_1\partial_1\varphi_2),$$

所以

$$\int_D \det(\nabla\varphi)dx = 0, \quad \forall\, \varphi \in W_0^{1,\infty}(\Omega).$$

于是有

$$\begin{aligned}&\text{mes}(D)^{-1}\int_D \det(A + \nabla\varphi(x))dx \\ &= \text{mes}(D)^{-1}\int_D [\det(A) + \det(\nabla\varphi(x)) + a_{11}\partial_2\varphi_2 + a_{22}\partial_1\varphi_1 \\ &\quad - a_{12}\partial_1\varphi_2 - a_{21}\partial_2\varphi_1]dx \\ &= \det(A).\end{aligned}$$

回忆第六讲中介绍过的判定弱极小点 u_0 的 Legendre-Hadamard 条件,

$$L_{p^j_\alpha p^k_\beta}(x, u_0(x), \nabla u_0(x))\pi^j_\alpha\pi^k_\beta \geqslant 0, \quad \forall\, \pi \in \mathbb{R}^{n\times N},\ \text{rank}(\pi) = 1,$$

它与凸性有关系. 事实上, 如果 $\forall\,(x,u) \in \Omega \times \mathbb{R}^N$, $p \to L(x,u,p)$ 是凸的, 并且 $L \in C^2$, 那么对于任意 u, Legendre-Hadamard 条件成立.

然而, Legendre-Hadamard 条件并不要求 L 对所有 $n\times N$ 矩阵是凸的, 只要求其对秩 1 的矩阵是凸的. 我们把满足条件 (以下省略 $(x,u)\in\Omega\times\mathbb{R}^N$)

$$L(\lambda B+(1-\lambda)C)\leqslant\lambda L(B)+(1-\lambda)L(C),$$
$$\forall\lambda\in[0,1],\forall B,C\in M^{n\times N},\quad \mathrm{rank}(B-C)=1$$

的 Lagrange 函数 L 称为秩 1 凸的.

各种凸性之间的关系如下:

$$凸\Longrightarrow 拟凸\Longrightarrow 秩\ 1\ 凸 \overset{当n=1或N=1时}{\Longrightarrow} 凸.$$

回到存在性问题, 借助于 Morrey-Acerbi-Fusco 定理, 我们有如下更一般的存在性定理.

定理 11.6 设 $L:\ M^{n\times N}\xrightarrow{C}\mathbb{R}^1$ 拟凸, 并存在常数 $C_2>C_1>0$ 使得

$$C_1|A|^p\leqslant L(A)\leqslant C_2(1+|A|^p),\quad 1<p<\infty,$$

又设 $v\in W^{1,p}(\Omega,\mathbb{R}^N)$, 则在 $E=W^{1,p}_v=\{u\in W^{1,p}(\Omega,\mathbb{R}^N)\big|u|_{\partial\Omega}=v|_{\partial\Omega}\}$ 中,

$$I(u)=\int_\Omega L\left(\nabla u(x)\right)dx$$

达到极小值.

习 题

1. 设 $\forall x\in\Omega$, $(u,p)\to L(x,u,p)$ 是凸的, 求证泛函

$$I(u)=\int_\Omega L(x,u(x),\nabla u(x))dx$$

对于 u 是凸的.

2. 在空间 $W^{1,q}(0,2\pi)(1\leqslant q<\infty)$ 中求序列 $\varphi_n(x)=\frac{1}{n}\sin nx$ 的弱极限.

3. 设 $\Omega\subset\mathbb{R}^n$ 是一个有界区域, 又设在空间 $W^{1,q}(\Omega)(1\leqslant q<\infty)$ 中, 序列 $\{u_j\}$ 弱收敛, 并且 $\{u_j\}$ 几乎处处收敛到一个函数 u_0. 问 $\{u_j\}$ 是否有子列在 $L^q(\Omega)$ 收敛? 收敛到什么? u_0 是否在 $W^{1,q}(\Omega)$ 中?

4. 设 $J=(0,1)$, $\lambda\in(0,1),\alpha,\beta\in\mathbb{R}^1$,

$$\varphi(x)=\begin{cases}\alpha,\ x\in(0,\lambda),\\ \beta,\ x\in(\lambda,1).\end{cases}$$

令 $\varphi_n(x)=\varphi(nx)$. 试证: $\varphi_n\rightharpoonup\lambda\alpha+(1-\lambda)\beta$.

5. 下列泛函是否在对应空间中弱下半连续?

(1) 设 Φ 是单变量的凸函数, c 是有界集 $\bar{\Omega}$ 上的连续函数, $I(u) = \int_\Omega(\Phi(|\nabla u(x)|) + c(x)|u(x)|^4)dx$, $W^{1,4}(\Omega)$.

(2) $I(u) = \int_\Omega[1 + |\nabla u(x)|^2]^{1/2}dx$, $W^{1,1}(\Omega)$.

(3) $I(u) = \int_{-1}^{1} t^2\dot{u}(t)^2dt$, $H^1(-1,1)$.

第十二讲　线性微分方程的边值问题与特征值问题

§12.1　线性边值问题与正交投影

在前几讲研究泛函极值存在性的方法中, 定义域的弱列紧性和泛函的序列弱下半连续性起着重要的作用.

然而正如第九讲提到的, 有一类对应于线性微分方程的特殊变分问题, 弱列紧性与序列弱下半连续性可以被 Hilbert 空间的完备性所取代, 只要用到 Hilbert 空间的正交投影就够了. 本讲介绍正交投影方法和它在变分问题中的应用.

回顾第十一讲中的 Poisson 方程 问题. 设 $\Omega \subset \mathbb{R}^n$ 是一个有界区域, 给定一个函数 $f \in L^2(\Omega)$, 求 $u : \Omega \to \mathbb{R}^1$ 满足下列方程:

$$\begin{cases} -\Delta u = f, & \Omega, \\ u = 0, & \partial\Omega. \end{cases}$$

在第十一讲中, 我们通过证明在空间 $H_0^1(\Omega)$ 中泛函

$$I(u) = \int_\Omega \left(\frac{1}{2} |\nabla u|^2 - fu \right) dx$$

的极小化序列有收敛子列来求得极值.

然而注意到 $H_0^1(\Omega)$ 是一个 Hilbert 空间, 它有内积

$$((u, v)) = \int_\Omega \nabla u \nabla v dx, \quad u, v \in H_0^1(\Omega),$$

泛函 I 的第一项正是这个空间的内积导出的模, 它的第二项也可以通过内积的形式表达出来. 事实上, 把

$$\varphi \mapsto \int_\Omega f \cdot \varphi dx$$

看成是 $H_0^1(\Omega)$ 上的线性泛函. 由 Schwarz 不等式得到

$$\left|\int_\Omega f \cdot \varphi dx\right| \leqslant \|f\|_2 \|\varphi\|_2,$$

利用 Poincaré不等式便有

$$\left|\int_\Omega f \cdot \varphi dx\right| \leqslant C \|f\|_2 \|\varphi\|_{H_0^1},$$

可见它还是连续的. 于是存在 $F \in (H_0^1(\Omega))^*$ 使得

$$F(\varphi) = \int_\Omega f \cdot \varphi dx.$$

按泛函分析中的 Riesz 定理, 这个连续线性泛函有内积表示, 即 $\exists\, u_0 \in H_0^1(\Omega)$, 使得

$$((u_0, \varphi)) = F(\varphi) = \int_\Omega f \varphi dx.$$

现在我们把 I 的 E-L 方程写成内积的形式:

$$((u, \varphi)) = \int_\Omega f \cdot \varphi dx = ((u_0, \varphi)), \quad \forall\, \varphi \in C_0^\infty(\Omega).$$

因为 $C_0^\infty(\Omega)$ 在 $H_0^1(\Omega)$ 中是稠密的, 所以 $u = u_0$ 就是解.

在这里好像既没有涉及极小化序列也没有用到序列弱下半连续性和弱列紧性, 就直接解出来了. 其实 Riesz 定理的证明过程本身, 就是在求解一个变分问题, 即, 在 Hilbert 空间中寻求一点到一个超平面的距离与垂足.

让我们回顾一下这个证明. 记

$$M = \{\eta \in H_0^1(\Omega) \mid F(\eta) = 0\},$$

它是 $H_0^1(\Omega)$ 的一个闭线性子空间. 如果 $\forall \varphi \in H_0^1(\Omega)$, 我们能够得到 φ 在 M 上的正交投影 η, 即 $\eta \in M$, 并且 $\varphi - \eta = \xi \perp M$, 那么我们只需取

$$v_0 = \begin{cases} F(\xi)\dfrac{\xi}{[\![\xi]\!]^2}, & \xi \neq 0, \\ 0, & \xi = 0, \end{cases}$$

就有 $((v_0,\varphi))=F(\varphi)$.

事实上, 若 $\xi=0$, 则上式显然成立. 若 $\xi\neq 0$, 那么 $\forall\varphi\in H_0^1(\Omega)$,

$$F(\varphi)=F(\xi),$$

而

$$((v_0,\varphi))=F(\xi)\left(\left(\frac{\xi}{[\![\xi]\!]^2},\varphi\right)\right)=F(\xi).$$

即得结论.

如果我们把抽象 Hilbert 空间中的内积具体化:

$$[\![u]\!]=\left(\int_\Omega |\nabla u|^2dx\right)^{1/2},$$

那么下列证明的过程正好就是求解变分问题的极小化过程.

现在回过来看正交投影的存在性 (图 12.1). 这可以先化归为一个求极值问题:

$$\min_{x\in M}[\![\varphi-x]\!],$$

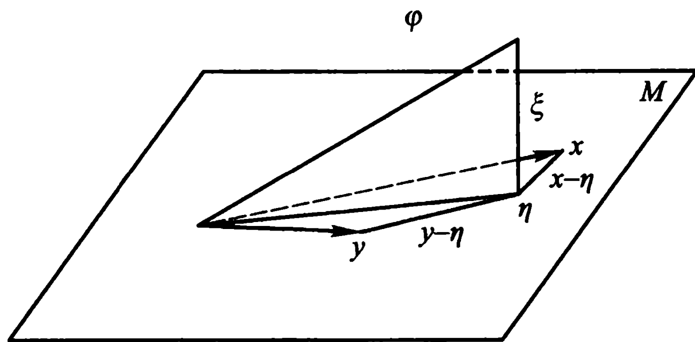


图 12.1

事实上, 如果存在 $\eta\in M$ 使得 $[\![\varphi-\eta]\!]=\min\limits_{x\in M}[\![\varphi-x]\!]$, 那么 $\xi=\varphi-\eta$ 满足

$$((\xi,x-\eta))=0,\quad \forall x\in M,$$

这是因为 $\forall x\in M,\forall t\in[0,1]$, 令

$$\begin{aligned}g(t)&=[\![\varphi-(tx+(1-t)\eta)]\!]^2\\&=[\![\varphi-\eta]\!]^2-2t((\varphi-\eta,x-\eta))+t^2[\![x-\eta]\!]^2,\end{aligned}$$

g 是 t 的一个二次函数, 它在 $t=0$ 处达到极小值, 所以

$$g'(0)=-2((\varphi-\eta,x-\eta))=0, \tag{12.1}$$

即

$$((\xi, x-\eta))=0.$$

这个极小值为什么是存在的? 令 $m=\inf_{x\in M}[\![\varphi-x]\!]$. 取极小化序列 $\{\eta_j\}\subset M$, 使得

$$[\![\eta_j-\varphi]\!]<m+\frac{1}{j},\quad \forall j.$$

(图 12.2) 问题是: $\{\eta_j\}$ 是否收敛? $\forall\varepsilon>0$, 利用平行四边形法则, 再注意到 M 是一个线性子空间, 我们有

$$\begin{aligned}[\![\eta_j-\eta_k]\!]^2&=2([\![\eta_j-\varphi]\!]^2+[\![\eta_k-\varphi]\!]^2)-4\left[\!\!\left[\frac{\eta_j+\eta_k}{2}-\varphi\right]\!\!\right]^2\\&\leqslant 4(m+\varepsilon)^2-4m^2,\quad 当\ j,k\ 充分大.\end{aligned}\tag{12.2}$$

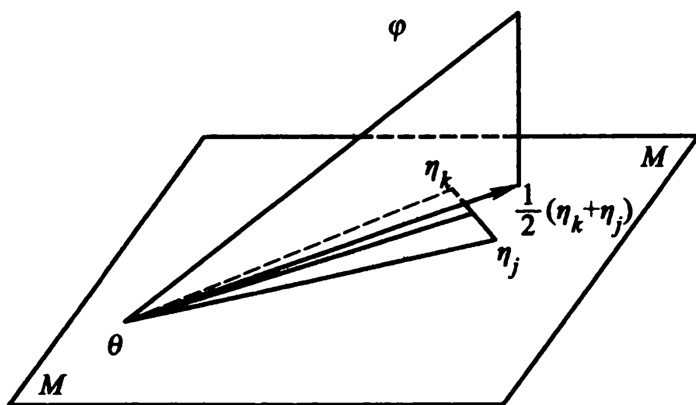


图 12.2

可见 $\{\eta_j\}$ 是一个 Cauchy 序列. 因为空间 $H_0^1(\Omega)$ 是完备的, 所以 $\{\eta_j\}$ 是收敛的, $\eta_j\to\eta$. 又因为 M 是闭的, 所以 $\eta\in M$, 并实现

$$[\![\eta-\varphi]\!]=\min_{x\in M}[\![\varphi-x]\!].$$

这是从另一个角度来验证 Dirichlet 原理. 由于问题的特殊性, 没有必要引入弱收敛, 然而空间的完备性起着关键的作用.

这个方法同样可以应用于求解变分不等式 (参看第六讲), 例如, 薄膜的障碍问题: 在有界区域 $\Omega\subset\mathbb{R}^n$ 上给定一个可测函数 $\psi(x)$ 作为障碍, 以及作为外力的函数 $f\in L^2(\Omega)$. 求薄膜 $u:\Omega\to\mathbb{R}^1$ 的平衡位置.

我们把这个问题提成一个变分问题. 取 Hilbert 空间 $H_0^1(\Omega)$ 上的一个闭凸集 $C=\{u\in H_0^1(\Omega)\,|\,u(x)\leqslant\psi(x)\text{ a.e.}\}$. 记 $(f,v)=\int_\Omega f\cdot v dx$. 求解 $u\in C$, 使得

$$((u,v-u))-(f,v-u)\geqslant 0,\quad \forall v\in C.\tag{12.3}$$

经上面的讨论, 存在 $u_0 \in H_0^1(\Omega)$ 使得

$$((u_0, v)) = (f, v), \quad \forall v \in H_0^1(\Omega).$$

于是化归为

$$((u - u_0, v - u)) \geqslant 0, \quad \forall v \in C. \tag{12.4}$$

根据前面的推理, 化归求

$$\min_{v \in C} [\![u_0 - v]\!] \tag{12.5}$$

事实上, 若 u 是这问题的解, 则因 $\forall v \in C,\ tv + (1-t)u \in C$, (12.1) 变为

$$g'(0) \geqslant 0,$$

即 (12.4). 至于 (12.5) 的解的存在性仍是通过极小化序列 $\{\eta_j\} \subset C$ 来得到的. 因为 C 是凸集, 所以在 (12.2) 中 $\frac{1}{2}(\eta_j + \eta_k) \in C$. 再由平行四边形等式, $\{\eta_j\}$ 还是一个 Cauchy 列, 而 C 是闭的, 所以在 C 中, 存在极限实现极小. 我们已经知道问题 (12.5) 的解就是 (12.3) 的解.

§12.2 特征值问题

设 $\Omega \subset \mathbb{R}^n$ 是一个区域, 对于哪些 $\lambda \in \mathbb{R}^1$, 边值问题

$$\begin{cases} -\Delta u = \lambda u, & \Omega, \\ u = 0, & \partial\Omega \end{cases} \tag{12.6}$$

有非零解 u? 和矩阵的特征值问题类似, 我们把它称为微分算子 $-\Delta$ 的特征值问题.

对应着非零解的数 λ 称为 “谱”. 如果这个非零解 u 是 $L^2(\Omega)$ 函数, 那么我们称它为特征函数, 称相应的 λ 为特征值. 在几何、力学、物理以及许多数学分支中经常遇到寻求这类特征值的问题.

特征值问题可以用约束变分方法处理. 正如第六讲所述, 令

$$I(u) = \int_\Omega |\nabla u|^2 \, dx,$$

以及

$$N(u) = \int_\Omega |u|^2 \, dx.$$

求

$$\min\{I(u) \mid u \in H_0^1(\Omega) \cap N^{-1}(1)\}. \tag{12.7}$$

如果 $\varphi_1 \in C^2(\bar{\Omega})$ 实现这个极小值, 那么应用 Lagrange 乘子法, 对于调节后的 Lagrange 函数的 E-L 方程, 存在 $\lambda_1 \in \mathbb{R}^1$, 满足

$$-\Delta\varphi_1 = \lambda_1\varphi_1 \text{ 于 } \Omega. \tag{12.8}$$

因为 $N(\varphi_1) = \int_\Omega |\varphi_1|^2\, dx = 1$, 所以 φ_1 是非零的. 再在 (12.8) 式的两边同乘以 φ_1 后积分, 得到

$$I(\varphi_1) = \int_\Omega |\nabla\varphi_1|^2\, dx = \lambda_1.$$

这表明

$$\lambda_1 = \min\{I(u) \mid u \in H_0^1(\Omega) \cap N^{-1}(1)\}. \tag{12.9}$$

下面我们来验证极小值问题 (12.7) 解的存在性. 为此假设 $\Omega \subset \mathbb{R}^n$ 是一个有界区域. 我们已经知道 I 是强制的并且是序列弱下半连续的, 剩下来要验证集合

$$M_1 = \{u \in H_0^1(\Omega) \mid N(u) = 1\}$$

是序列弱闭的; 即若 $\{u_j\} \subset M_1$, $u_j \rightharpoonup u$ $(H_0^1(\Omega))$, 则 $u \in M_1$.

按 Rellich-Kondrachov 紧性定理, 对于有界区域 Ω, $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ 的嵌入是紧的. 由假设 $u_j \rightharpoonup u$ $H_0^1(\Omega)$, 蕴含了 u_j 存在子列

$$u_j' \to u \quad (L^2(\Omega)).$$

又由

$$\int_\Omega u_j^2 dx = 1,$$

推得

$$\int_\Omega u^2 dx = 1,$$

即 $u \in M_1$, 亦即 M_1 是弱闭的.

求得第一个特征值以后, 类似于矩阵特征值的问题, 我们自然要问: 还有没有其他的特征值了? 怎样求得它们?

我们继续使用约束极值的思想, 取集合

$$M_2 = \left\{u \in M_1 \,\middle|\, \int_\Omega u\varphi_1 dx = 0\right\},$$

求解

$$\min\{I(u) \mid u \in M_2\}.$$

如果 $\varphi_2 \in M_2$ 是它的一个极小点, 那么同理 $\varphi_2 \neq 0$, 并且存在 Lagrange 乘子 λ_2, μ_2 使得

$$-\Delta\varphi_2 = \lambda_2\varphi_2 + \mu_2\varphi_1. \tag{12.10}$$

我们要证明: $\mu_2 = 0$. 事实上, 对 (12.8) 两边同乘以 φ_2 然后积分得到

$$\int_\Omega \nabla\varphi_2\nabla\varphi_1 dx = \int_\Omega \nabla\varphi_1\nabla\varphi_2 dx = \lambda_1\int_\Omega \varphi_1\varphi_2 dx = 0;$$

在等式 (12.10) 的两边同乘以 φ_1 后积分, 得到

$$\int_\Omega \nabla\varphi_2\nabla\varphi_1 dx = \mu_2\int_\Omega |\varphi_1|^2 dx.$$

所以 $\mu_2 = 0$.

这样一来就有

$$-\Delta\varphi_2 = \lambda_2\varphi_2 \text{ 于 } \Omega.$$

也就证明了 λ_2 是一个特征值, φ_2 是相应的特征函数, $\varphi_2 \neq \varphi_1$. 还有

$$\lambda_2 = I(\varphi_2) = \int_\Omega |\nabla\varphi_2|^2\, dx = \min\{I(u) \mid u \in M_2\} \geqslant \lambda_1.$$

为了证明 I 在 M_2 上能达到极小值, 我们还要验证 M_2 的弱闭性. 设 $u_j \rightharpoonup u\ (H_0^1(\Omega))$, 则由 $\{u_j\} \subset M_1$ 以及 M_1 是弱闭的, 推出 $u \in M_1$, 再由 $\int_\Omega u_j\varphi_1 dx = 0$, 推出

$$\int_\Omega u\varphi_1 dx = 0.$$

即 $u \in M_2$, 亦即 M_2 是弱闭的.

如此继续下去, 令

$$M_n = \left\{u \in M_{n-1}\middle| \int_\Omega u\varphi_{n-1} dx = 0\right\},$$

同理证明它还是弱闭的, 从而极值问题

$$\min\{I(u) \mid u \in M_n\}$$

有解 $\varphi_n \neq 0$, 满足

$$-\Delta\varphi_n = \lambda_n\varphi_n + \sum_{j=1}^{n-1}\mu_j\varphi_j.$$

同理用数学归纳法可证: $\mu_1 = \cdots = \mu_{n-1} = 0$, 便得到

$$-\Delta\varphi_n = \lambda_n\varphi_n,$$

其中

$$\lambda_n = I(\varphi_n) = \min\{I(u) \mid u \in M_n\} \geqslant \lambda_{n-1}.$$

于是我们有如下定理.

定理 12.1 设 $\Omega \subset \mathbb{R}^n$ 是一个有界区域, 则方程 (12.6) 存在一列特征值: $0 < \lambda_1 \leqslant \lambda_2 \leqslant \cdots \to +\infty$, 以及对应的特征函数 $\{\varphi_1, \varphi_2, \cdots\} \subset H_0^1(\Omega)$, 满足

$$\begin{cases} \Delta\varphi_i = \lambda_i\varphi_i, \\ \displaystyle\int_\Omega |\varphi_i|^2\, dx = 1, \end{cases} \quad i = 1, 2, \cdots$$

以及

$$((\varphi_i, \varphi_j)) = \int_\Omega \nabla\varphi_i\nabla\varphi_j dx = \lambda_i\int_\Omega \varphi_i\varphi_j dx = 0, \quad \forall i \neq j.$$

证明 现在只需要再证明: $\lambda_n \to +\infty$.

我们用反证法, 如果 $\exists C > 0$ 使得 $\lambda_n \leqslant C$, 那么

$$\int_\Omega |\nabla\varphi_j|^2\, dx = \lambda_j\int_\Omega |\varphi_j|^2\, dx = \lambda_j \leqslant C, \quad \forall j.$$

这表明 $\{\varphi_j\}_1^\infty$ 是 $H_0^1(\Omega)$ 中的有界列, 从而有弱收敛子列

$$\varphi_{j'} \rightharpoonup \varphi^*(H_0^1(\Omega)).$$

一方面, 应用 Rellich-Kondrachov 紧性定理, 有

$$\varphi_{j'} \to \varphi^*(L^2(\Omega)). \tag{12.11}$$

另一方面, 由 $\int_\Omega \varphi_i\varphi_j dx = 0,\ i \neq j$, 可见

$$\int_\Omega |\varphi_i - \varphi_j|^2 dx = \int_\Omega (\varphi_i^2 + \varphi_j^2) dx = 2, \quad i \neq j. \tag{12.12}$$

将 (12.11) 代入 (12.12) 得到

$$\int_\Omega |\varphi^* - \varphi_j|^2 dx = 2, \quad \forall j.$$

特别地, 取 $j = j'$ 便有 $\int_\Omega |\varphi^* - \varphi_j'|^2 dx = 2$. 这显然与 (12.11) 矛盾. □

注 12.1 由椭圆型方程的正则性理论, 特征函数 $\varphi_1, \varphi_2, \cdots$ 在区域 Ω 内实际上都是无穷次连续可微的. 如果区域 Ω 的边界还是光滑的, 那么它们在 $\bar{\Omega}$ 上也是无穷次连续可微的. □

§12.3 特征展开

我们已经证明了特征函数组 $\{\varphi_i\}_1^\infty$ 不仅在 $H_0^1(\Omega)$, 而且在 L^2 空间的内积下也是两两正交的, 从而形成 $L^2(\Omega)$ 空间中的一组正交基. 现在还要证明对于 $L^2(\Omega)$ 空间来说, 这组正交基是完备的.

$\forall u \in L^2(\Omega)$, 令

$$c_n = \int_\Omega u(x)\varphi_n(x)dx, \quad \forall n,$$

称其为 u 的广义 Fourier 系数. 作部分和

$$s_m(x) = \sum_{n=1}^m c_n\varphi_n(x).$$

所谓 $\{\varphi_i\}_1^\infty$ 是完备的, 是指当 $m \to \infty$ 时,

$$s_m(x) \to u(x) \quad (L^2(\Omega)).$$

事实上, 一方面由正交性,

$$\begin{aligned}\int_\Omega |\nabla(u-s_m)|^2 dx &= \int_\Omega |\nabla u|^2 dx - 2\int_\Omega \nabla u \nabla s_m dx + \int_\Omega |\nabla s_m|^2 dx \\ &= \|\nabla u\|_2^2 - 2\sum_{n=1}^m \lambda_n |c_n|^2 + \sum_{n=1}^m \lambda_n |c_n|^2 \qquad (12.13) \\ &= \|\nabla u\|_2^2 - \sum_{n=1}^m \lambda_n |c_n|^2,\end{aligned}$$

推得

$$\sum_{n=1}^m \lambda_n |c_n|^2 \leqslant \int_\Omega |\nabla u|^2 dx. \qquad (12.14)$$

另一方面, 因为 $u - s_m \in M_{m+1}$, 所以

$$\int_\Omega |\nabla(u-s_m)|^2 dx \geqslant \lambda_{m+1}\int_\Omega |u-s_m|^2 dx.$$

利用 (12.13) 推出

$$\|u-s_m\|_2^2 \leqslant \frac{1}{\lambda_{m+1}}\|\nabla u\|_2^2 \to 0 \quad (m\to\infty).$$

于是有 Fourier 展开

$$u(x) = \sum_{n=1}^\infty c_n\varphi_n(x) \quad (\text{在 } L^2 \text{ 平均意义下收敛})$$

以及 Parseval 等式

$$\int_\Omega |u|^2 dx = \lim_{m\to\infty} \|s_m\|_2^2 = \lim_{m\to\infty} \sum_{i,j=1}^m c_i c_j \int_\Omega \varphi_i \varphi_j dx = \sum_{m=1}^\infty |c_m|^2.$$

此外, 我们还能进一步有

(1) $s_m \to u$ 于 $H_0^1(\Omega)$,

(2) $\int_\Omega |\nabla u|^2 dx = \sum_{n=1}^\infty \lambda_n |c_n|^2$.

事实上, 联合 (12.13) 与 (12.14), 再令 $m > n \to \infty$, 就得到

$$\|\nabla(s_m - s_n)\|_2^2 = \sum_{j=n+1}^m \lambda_j |c_j|^2 \to 0.$$

于是有 $s_m \to u\,(H_0^1(\Omega))$. 这时

$$\|\nabla u\|_2^2 = \lim_{m\to\infty} \|\nabla s_m\|_2^2 = \sum_{j=1}^\infty \lambda_j |c_j|^2.$$

注 12.2 在 $n = 1$ 的特殊情形, 我们考虑如下 Sturm-Liouville 问题. 在区间 $J = [a, b]$ 上给定 $p \in C^1(J)$, $q \in C(J)$, 假设存在常数 $\alpha > 0$ 使得

$$p(x) \geqslant \alpha.$$

又设 $q(x) \geqslant 0, \forall x \in J$, 我们把

$$Lu = -(pu')' + qu$$

称为 Sturm-Liouville 算子. □

考察下列特征值问题:

$$\begin{cases} Lu = \lambda u & \text{于 } J, \\ u(a) = u(b) = 0. \end{cases}$$

和前面的讨论完全一样, 在空间 $H_0^1(J)$ 上, 考察泛函

$$I(u) = \int_J \frac{1}{2}(p(t)|u'(t)|^2 + q(t)|u(t)|^2)dt,$$

不过现在我们针对泛函 I 定义 $H_0^1(J)$ 的一个等价模与内积如下:

$$\|u\| = \left(\int_J (p(t)|u'(t)|^2 + q(t)|u(t)|^2)dt\right)^{1/2},$$

$$((u, v)) = \int_J (p(t)u'(t)v'(t) + q(t)u(t)v(t))dt.$$

通过逐次引入约束 $M_1 = \{u \in H_0^1(J) \mid \int_J u^2 dt = 1\}, M_2, \cdots$, 我们得到特征值与特征函数

$$0 < \lambda_1 \leqslant \lambda_2 \leqslant \cdots \to +\infty,$$

以及

$$\varphi_1, \varphi_2, \cdots,$$

我们不难验证 $\varphi_n \in C^2(J)$, $\forall n$.

它们满足

$$\int_J \varphi_n \varphi_m dx = \delta_{nm}, \quad \forall n, m,$$

以及

$$L\varphi_n = \lambda_n \varphi_n, \quad \forall n.$$

此外, $\forall u \in H_0^1(J)$, 有广义 Fourier 展开

$$u(x) = \sum_{n=1}^{\infty} c_n \varphi_n(x)$$

在 $H_0^1(\Omega)$ 意义下收敛, 其中

$$c_n = \int_J u(x)\varphi_n(x)dx.$$

例 12.1 设 $p = 1, q = 0, J = [0, \pi]$, 则

$$Lu = -\ddot{u}.$$

这时

$$\lambda_n = n^2, \quad \varphi_n(x) = \sin nx, \ n = 1, 2, \cdots. \qquad \Box$$

注 12.3 除了 Dirichlet 问题外, 我们还可以讨论其他类型边值的特征值问题, 例如 Neumann 问题:

$$\begin{cases} -(pu')' + qu = \lambda u & 于 (a, b), \\ u'(a) = u'(b) = 0. \end{cases} \qquad \Box$$

正如在第一讲与第六讲中所指出的, 对于 Neumann 问题我们用空间 $H^1(\Omega)$ 取代空间 $H_0^1(\Omega)$. 这是因为用分部积分公式

$$\int_a^b (pu')'\varphi dx = (pu')\varphi|_a^b - \int_a^b (pu')\varphi' dx$$

于积分形式的 E-L 方程

$$\int_a^b (pu'\varphi' + (q-\lambda)u\varphi)dx = 0, \quad \forall \varphi \in H^1(J),$$

同时导出微分方程

$$-(pu')' + qu = \lambda u$$

和边值条件

$$u'(a) = u'(b) = 0.$$

然而在空间 $H^1(J)$ 上, 泛函

$$I(u) = \int_a^b (p|u'|^2 + q|u|^2)dx$$

没有强制性. 我们用它的一个子空间

$$X = \left\{ u \in H^1(J) \,\middle|\, \int_a^b u dx = 0 \right\},$$

即, 与常值函数正交的子空间来取代 $H_0^1(J)$.

在子空间 X 上 Poincaré 不等式还是成立的. 证明方法和原来的一样, 因为 $\int_a^b u dx = 0$, 所以 $\exists \xi \in (a,b)$ 使得 $u(\xi) = 0$, 于是

$$|u(x)| = \left| \int_\xi^x u'(t)dt \right| \leqslant (b-a)^{1/2} \left(\int_a^b |u'|^2 dt \right)^{1/2}$$

我们现在可以在这个子空间 X 上求泛函 I 的极小值, 换句话说, 相当于在空间 $H^1(J)$ 中添加一个积分约束条件

$$\int_J u(t)dt = 0.$$

值得注意的是, 这个约束所产生的 Lagrange 乘子在方程中自然消失.

注意到当 $q = 0$ 时, 非零常数本身就是对应于特征值为 0 的一个特征函数. 此时在 Neumann 边值条件下, Sturm-Liouville 问题的特征值为 $\{0, \lambda_1, \lambda_2, \cdots\}$, 特征函数为 $\{1/\sqrt{b-a}, \varphi_1, \varphi_2, \cdots\}$.

对于 Laplace 算子的 Neumann 边值条件下的特征值问题:

$$\begin{cases} \Delta u = \lambda u & \text{于 } \Omega, \\ \dfrac{\partial u}{\partial n}\Big|_{\partial\Omega} = 0, \end{cases}$$

其中 n 是 $\partial\Omega$ 的单位法方向, 我们也引进子空间

$$X = \left\{ u \in H^1(\Omega) \,\middle|\, \int_\Omega u(x)dx = 0 \right\}.$$

把 Poincaré 不等式推广到这个子空间, 再在这个子空间上 (添加一个约束条件) 求泛函 $I(u) = \int_\Omega |\nabla u|^2 dx$ 的极小值.

注 12.4 同样的方法可以用来研究一个无边的紧 Riemann 流形 (M, g) 上 Laplace-Beltrami 算子的特征值问题:

$$\Delta_g u = \frac{1}{\sqrt{g}} \sum_{ij} \partial_i (g^{ij} \sqrt{g} \partial_j u),$$

其中 $g = \det(g_{ij})$, 没有边值条件. □

特别值得注意的是周期边值条件的特征值问题:

$$\begin{cases} -(pu')' + qu = \lambda u & 于 (a, b), \\ u(a) = u(b), \quad u'(a) = u'(b), \end{cases}$$

它可以看成是无边界流形 S^1 上的特征值问题. 这时我们用周期函数空间 $H^1_{\text{per}}(J)$ 取代 $H^1_0(J)$, 并用 Wirtinger 不等式 (见第十三讲引理 13.4) 取代 Poincaré 不等式来验证泛函的强制性.

注 12.5 以上确定特征值与特征函数的步骤正是 $\mathbb{R}^n$ 中二次曲面确定标准型的几何方法, $\lambda_1^{-1}, \lambda_2^{-1}, \cdots$ 是逐次的长轴. □

§12.4 特征值的极小极大刻画

前面关于特征值 λ_n 的极值刻画是递归式的, 也就是说必须逐次确定了前 $n-1$ 个特征函数 $\varphi_1, \cdots, \varphi_{n-1}$ 之后才能确定 λ_n. 下面的极小极大定理使得 λ_n 的确定可以一步到位.

定理 12.2 (Courant 极小极大定理)

$$\lambda_n = \max_{E_{n-1}} \min_{u \in E_{n-1}^{\perp} \setminus \{\theta\}} \frac{\displaystyle\int_\Omega |\nabla u|^2 dx}{\displaystyle\int_\Omega |u|^2 dx},$$

其中 E_{n-1} 是 $H^1_0(\Omega)$ 中的任意 $n-1$ 维线性子空间.

证明 设 $v_1, \cdots, v_{n-1} \in H_0^1(\Omega)$ 是一组线性无关的函数, 令

$$E_{n-1} = \mathrm{span}\{v_1, v_2, \cdots, v_{n-1}\},$$

以及

$$\mu(E_{n-1}) = \min_{u \in E_{n-1}^{\perp} \setminus \{\theta\}} \frac{\displaystyle\int_\Omega |\nabla u|^2 dx}{\displaystyle\int_\Omega |u|^2 dx}.$$

一方面, 我们证明:

$$\mu(E_{n-1}) \leqslant \lambda_n.$$

事实上, 令 $\{\varphi_1, \cdots, \varphi_n\}$ 为前 n 个特征函数, 则 $(E_{n-1}^{\perp} \setminus \{\theta\}) \cap \mathrm{span}\{\varphi_1, \cdots, \varphi_n\} \neq \varnothing$, 即有非零的 u 满足:

$$\begin{cases} u = \displaystyle\sum_{i=1}^{n} c_i \varphi_i, \\ \displaystyle\int_\Omega u v_j dx = 0, \end{cases} \quad j = 1, \cdots, n-1,$$

或

$$\sum_{i=1}^{n} c_i \int_\Omega \varphi_i v_j dx = 0, \quad j = 1, \cdots, n-1.$$

这个以 $\left(\int_\Omega \varphi_i v_j dx\right)$ 为系数的 n 个未知数 $c_1, \cdots, c_n$ 的 $n-1$ 个方程组必有非平凡解. 再由 $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$, 推出

$$\mu(E_{n-1}) \leqslant \frac{\displaystyle\int_\Omega |\nabla u|^2 dx}{\displaystyle\int_\Omega |u|^2 dx} = \frac{\displaystyle\sum_{i=1}^{n} \lambda_i |c_i|^2}{\displaystyle\sum_{i=1}^{n} |c_i|^2} \leqslant \lambda_n.$$

另一方面, 特别地取 $\tilde{E}_{n-1} = \{\varphi_1, \cdots, \varphi_{n-1}\}$, 便有

$$\max_{E_{n-1}} \mu(E_{n-1}) \geqslant \mu(\tilde{E}_{n-1}) = \lambda_n.$$

即得结论. □

习 题

1. 设 $\Omega \subset \mathbb{R}^n$ 是一个有界区域, 又设 $r \in C(\bar{\Omega})$ 是一个正连续函数. 试证:

$$-\Delta u(x) = \lambda r(x)u(x), \quad x \in \Omega,$$
$$u(x) = 0, \quad x \in \partial\Omega$$

存在无穷多特征值与特征函数. 它们依下列定义的内积两两正交:

$$(u, v) = \int_\Omega u(x)v(x)r(x)dx.$$

2. 沿用上题记号. 设 $a_{ij} \in C^1(\bar{\Omega}), c \in C(\bar{\Omega})$, 又设存在 $\alpha > 0$, 使得

$$\sum_{i,j=1}^n a_{i,j}(x)\xi_i\xi_j \geqslant \alpha|\xi|^2, \ \forall\, \xi = (\xi_1, \cdots, \xi_n) \in \mathbb{R}^n.$$

试证:

$$-\sum_{i,j=1}^n \partial_i(a_{i,j}(x)\partial_j u(x)) + c(x)u(x) = \lambda u(x), \quad x \in \Omega,$$
$$u(x) = 0, \quad x \in \partial\Omega$$

存在无穷多特征值与特征函数.

3. 设 H 是一个 Hilbert 空间, $a : H \times H \to \mathbb{R}^1$ 是一个双线性有界泛函, $V \subset H$ 是一个线性子空间. 又设存在 $c > 0$ 使得

$$a(v, v) \geqslant c\|v\|^2, \ \forall\, v \in V,$$

则 $(V, a(\cdot, \cdot))$ 是一个 Hilbert 空间.

试证: $\forall\, F \in V^*$, 存在唯一 $u \in V$, 使得

$$a(u, v) = F(v), \ \forall\, v \in V.$$

4. 在题 2 中, 设 $f \in L^2(\Omega)$, 试证:

$$-\sum_{i,j=1}^n \partial_i(a_{i,j}(x)\partial_j u(x)) + c(x)u(x) = f(x), \quad x \in \Omega,$$
$$u(x) = 0, \quad x \in \partial\Omega$$

存在唯一解 $u \in H_0^1(\Omega)$.

第十三讲 存在性与正则性

在上两讲, 我们把一个 E-L 方程解的存在性化归为寻求泛函的极值. 在一定条件下, 极小点 u 是在某个 Sobolev 空间的 $W^{1,q}(\Omega)$ 中通过极小化序列方法得到的, 只能在推广的意义下满足方程

$$\int_\Omega [L_u(x,u(x),\nabla u(x))\varphi(x)+L_p(x,u(x),\nabla u(x))\nabla\varphi(x)]dx=0,\quad \forall\varphi\in C_0^\infty(\Omega)$$

即这个极小点是对应的 E-L 方程的广义解.

然而对于含有一阶导数的泛函, 其 E-L 方程是一个二阶微分方程, 只有当解是二次可微的函数时才能成为通常意义下的微分方程的解. 从微分方程的角度看, 还需要回答这个广义解有没有足够的可微性来满足对应的 E-L 微分方程? 也就是说, 能否从所得的广义解 u 导出 $u\in C^2$? 或更高的可微性? 甚至解析性? 这就是所谓的正则性问题.

我们把寻找广义解的问题称为 “存在性” 问题, 而把广义解的可微性问题称为 “正则性” 问题.

由此可见, 在直接方法中 “存在性” 问题与 “正则性” 问题是一对孪生的问题.

在 Hilbert 的 23 个问题中, 第 19 问题是关于正则变分问题 (实际上是关于椭圆型方程) 解的解析性的, 第 20 问题是关于正则变分边值问题可解性 (即, 存在性) 的.

§13.1 正则性 $(n=1)$

以下总假设 Lagrange 函数 $L \in C^2$, 问什么时候泛函 I 的极小点 $u^* \in C^2(J)$?

引理 13.1 设 $u^* \in C^1(J)$ 是 $I(u) = \int_J L(t, u(t), \dot{u}(t))dt$ 的一个极小点, 如果 $\det(L_{p_ip_j}(t, u^*(t), \dot{u}^*(t))) \neq 0$, $\forall t \in J$, 那么 $u^* \in C^2(J)$.

证明 记

$$q(t) = \int_a^t L_u(t, u^*(t), \dot{u}^*(t))dt - c,$$

其中 c 是一个常数. 利用积分形式的 E-L 方程，我们有

$$L_p(t, u^*(t), \dot{u}^*(t)) = q(t).$$

令 $\varphi : \bar{J} \times \mathbb{R}^N \to \mathbb{R}^N$ 为

$$\varphi(t, p) = L_p(t, u^*(t), p) - q(t).$$

我们知道 $\varphi \in C^1(\bar{J} \times \mathbb{R}^N, \mathbb{R}^N)$ 满足

$$\det(\varphi_p(t, u^*(t))) = \det(L_{pp}(t, u^*(t), \dot{u}^*(t))) \neq 0,$$

以及

$$\varphi(t, \dot{u}^*(t)) = 0.$$

由隐函数定理, 方程

$$\varphi(t, p) = 0$$

有唯一的局部 C^1 解, 即 $\forall t_0 \in \bar{J}, \exists$ 邻域 $U = U(t_0)$ 以及唯一的 $\lambda \in C^1(U, \mathbb{R}^N)$, 使得在 $(t, \dot{u}^*(t))$ 的邻域内满足 $\varphi(t, \lambda(t)) = 0$.

于是 $\dot{u}^*(t) = \lambda(t) \in C^1$, 即得 $u^* \in C^2$. □

条件 $\det(L_{p_ip_j}(t, u^*(t), \dot{u}^*(t))) \neq 0$ 是不可缺的, 当它不满足时, 存在这样一个泛函, 其极值函数是 C^1 的, 但不是 C^2 的.

例 13.1 设 $L(t, u, p) = u^2(p - 2t)^2$, 以及 $M = \{u \in C^1([-1, 1]) \,|\, u(-1) = 0, u(1) = 1\}$, 则泛函

$$I(u) = \int_{-1}^1 u^2(\dot{u}(t) - 2t)^2 dt$$

有极小点

$$u^*(t)=\begin{cases}0, & t<0,\\ t^2, & t\geqslant 0.\end{cases}$$

易见 $u^*\in C^1\backslash C^2([-1,1])$. 这时 $L_{pp}(t,u^*(t),\dot u^*(t))=2(u^*(t))^2=0$, 当 $t<0$. □

以上引理用隐函数定理把已知的 C^1 解抬高到 C^2, 但还不足以保证解的正则性. 因为在第九讲中我们曾经指出过, C^1 空间不是适用于直接方法的函数空间, 在直接方法中我们得到的解只能在某个 Sobolev 空间中, 所以为了得到解的正则性, 我们还需要证明所得的广义解实际上是 C^1 解.

定理 13.1 当 $1<r<\infty$ 时, 设 L 满足如下增长条件:

$$|L(t,u,p)|+|L_u(t,u,p)|+|L_p(t,u,p)|\leqslant C(1+|p|^r),$$

而当 $r=\infty$ 时, 对 L 不加增长条件.

又设矩阵 $(L_{p_ip_j}(t,u,p))$, $\forall(t,u,p)\in\bar J\times\mathbb{R}^N\times\mathbb{R}^N$ 都是正定的.

若 $u^*\in W^{1,r}(J,\mathbb{R}^N)$ 是泛函 $I(u)=\int_\Omega L(t,u(t),\dot u(t))dt$ 的一个极小点, 则改变 u^* 在一个零测集上的值以后, $u^*\in C^2$.

证明 由引理 13.1, 只要能证明 $u^*\in C^1$ 就够了.

定义函数 $\varphi: J\times\mathbb{R}^N\times\mathbb{R}^N\times\mathbb{R}^N\to\mathbb{R}^1$,

$$\varphi(t,u,p,q)=L_p(t,u,p)-q.$$

由假设 $\det(L_{p_ip_j}(t,u,p))\neq 0$, $\forall(t,u,p)\in\bar J\times\mathbb{R}^N\times\mathbb{R}^N$, 以及隐函数定理, 存在唯一的局部 C^1 解

$$p=\lambda(t,u,q). \tag{13.1}$$

一方面, 我们指出这个解是整体唯一的. 事实上, 倘若 p_1,p_2 都满足 (13.1), 那么 $q=L_p(t,u,p_1)=L_p(t,u,p_2)$. 于是

$$0=(L_p(t,u,p_1)-L_p(t,u,p_2),p_1-p_2)=(B(p_1-p_2),p_1-p_2),$$

其中

$$B=\int_0^1 L_{pp}(t,u,p_1+\tau(p_2-p_1))d\tau.$$

由假设, B 是正定的, 得到 $p_1=p_2$.

另一方面, 因为 $u^* \in W^{1,r}(J, \mathbb{R}^N)$, 所以当 $1 < r < \infty$ 时, $L_u(t, u^*(t), \dot{u}^*(t)) \in L^1(J)$, 而当 $r = \infty$ 时, $L_u(t, u^*(t), \dot{u}^*(t)) \in L^\infty(J)$. 总之,

$$q(t) = \int_{t_0}^{t} L_u(s, u^*(s), \dot{u}^*(s))ds - c \text{ 是绝对连续的.}$$

由积分形式的 E-L 方程

$$\int_{t_0}^{t} L_u(t, u^*(t), \dot{u}^*(t))dt - L_p(t, u^*(t), \dot{u}^*(t)) = c, \quad \text{a.e. } t \in J,$$

可见

$$q(t) = L_p(t, u^*(t), \dot{u}^*(t)), \quad \text{a.e. } t \in J,$$

以及

$$\dot{u}^*(t) = \lambda(t, u^*(t), q(t)), \quad \text{a.e. } t \in J.$$

我们已经知道 $q(t)$ 是绝对连续的, 代入上式, 在改变一个零测集上的值后, 得到 $\dot{u}^*(t)$ 是连续的, 即得 $u^* \in C^1$, 再由引理 13.1 就得到 $u^* \in C^2$.

在上面的证明中, 矩阵

$$(L_{p_i p_j}(t, u, p)), \ \forall (t, u, p) \in \bar{J} \times \mathbb{R}^N \times \mathbb{R}^N$$

的整体正定性条件对于保证解的整体唯一性以及解的正则性是非常关键的. 这是因为 Sobolev 空间 $W^{1,r}$ 中的函数的导数可以是不连续的, 沿着这类函数的图像, 时间上相邻的两点在相空间中可能不在同一个邻域内. 而隐函数定理则是对相空间中的邻域来说的, 所以不再适用. □

下面是去掉整体的正定性假定后, 解不可微的一个例子.

例 13.2 设 $L(p) = (p^2 - 1)^2$, 以及 $M = \{u \in \text{Lip}([0, 1]) \mid u(0) = u(1) = 0\}$, 则泛函

$$I(u) = \int_0^1 (\dot{u}^2(t) - 1)^2 dt$$

有极小值 0.

这时 $L_{pp}(t, u, p) = 4(3p^2 - 1)$ 不是正定的.

如果 $u \in C^1$ 是极小点, 那么 $\dot{u}(t) = \pm 1$; 但无论 $\dot{u}(t) = 1$ 或者 $\dot{u}(t) = -1$, 都不可能有满足边值条件 $u(0) = u(1) = 0$ 的解, 也就是没有一个 C^1 函数实现其极小值. 相反, 有不可数个属于 Lip 类、满足 $\dot{u}(t) = \pm 1$ 的锯齿形函数都是这个变分问题的解 (见图 13.1). □

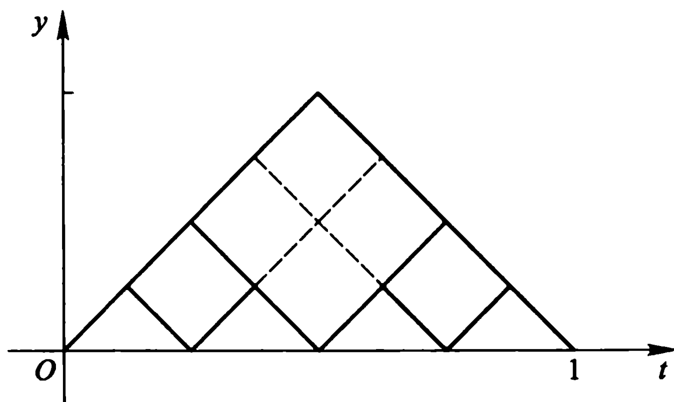


图 13.1 无穷多个锯齿形函数 $\dot{u}=\pm 1$ 都是解, 但没有 C^1 解

最后, 我们举例说明如果 Lagrange 函数不满足凸性, 那么泛函的序列弱下半连续性就未必成立.

例 13.3 (Bolza) 设 $L(u,p)=u^2+(p^2-1)^2$, 以及

$$M=W_0^{1,4}(0,1),$$

则泛函

$$I(u)=\int_0^1[u^2(t)+(\dot{u}(t)^2-1)^2]dt$$

的下确界为 0, 但在 M 中没有极小点.

事实上, 一方面, $I(u)\geqslant 0$; 另一方面, 定义一串锯齿形函数为极小化序列

$$u_j(t)=\begin{cases}t-\dfrac{k}{j}, & t\in\left[\dfrac{k}{j},\dfrac{2k+1}{2j}\right],\\ -t+\dfrac{k+1}{j}, & t\in\left[\dfrac{2k+1}{2j},\dfrac{k+1}{j}\right],\end{cases}$$

其中 $j=2,3,\cdots,0\leqslant k\leqslant j-1$.

由 $|\dot{u}_j(t)|=1$ a.e., 以及 $|u_j|\leqslant\frac{1}{2j}$, 可见

$$I(u_j)\leqslant\frac{1}{4j^2}\to 0.$$

从而 $\inf_{u\in M}I=0$.

我们指出: I 在 $W_0^{1,4}(0,1)$ 中没有极小值. 因若不然, 存在 $u_0\in W_0^{1,4}(0,1)$, 使得 $I(u_0)=0$, 则 $u_0(t)$=0 a.e., 从而 $\dot{u}_0(t)$=0 a.e., 如此则 $I(u_0)=1$, 矛盾 (见图 13.2). □

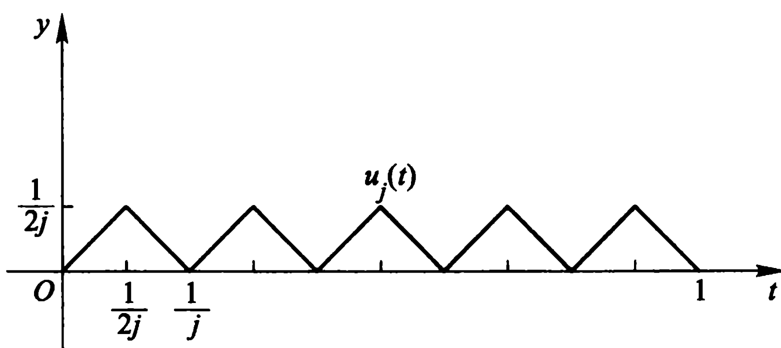


图 13.2 锯齿形极小化序列 u_j

§13.2 正则性续 $(n > 1)$

偏微分方程解的正则性问题要比常微分方程困难得多, 需要专门的学问. 为使读者有一个大致的了解, 仅举一例如下.

定理 13.2 (Weyl) 设 $u \in L^1_{\text{loc}}(\Omega)$ 满足

$$\int_\Omega u \cdot \Delta\varphi dx = 0, \ \forall \varphi \in C_0^\infty(\Omega), \tag{13.2}$$

则在修改一个零测集上的函数值后, $u \in C^\infty(\Omega)$.

我们把满足 (13.2) 的局部 L^1 函数称为弱调和函数. 证明的想法来自调和函数的平均值性质.

若 $u \in C^2(\Omega)$, 且 $\Delta u = 0$ 于 Ω, 则

$$u(x) = \frac{1}{r^{n-1}\varpi_n}\int_{\partial B_r(x)} u(y)d\sigma = \frac{n}{r^n\varpi_n}\int_{B_r(x)} u(y)dy, \ \forall B_r(x) \subset \Omega, \tag{13.3}$$

其中 ϖ_n 是 $\mathbb{R}^n$ 中单位球面 S 的面积. 我们把 (13.3) 称为函数 u 的平均值性质.

事实上, 设 $B_r(x) \subset \Omega$, 取 $\rho \in (0, r)$, 则

$$0 = \int_{B_\rho(x)} \Delta u(y)dy = \int_{\partial B_\rho(x)} \frac{\partial u}{\partial n} d\sigma = \rho^{n-1}\frac{\partial}{\partial\rho}\int_{|w|=1} u(x+\rho w)dw.$$

这表明

$$\int_{|w|=1} u(x+\rho w)dw = \text{const} = \varpi_n u(x),$$

所以

$$u(x) = \frac{1}{r^{n-1}\varpi_n}\int_{\partial B_r(x)} u(y)d\sigma = \frac{n}{r^n\varpi_n}\int_{B_r(x)} u(y)dy.$$

我们来推广并利用这个结论证明解的光滑性.

回顾第十讲中的光滑化算子的性质: $\forall\delta>0$, 令 $\Omega_\delta=\{x\in\Omega\,|\,d(x,\partial\Omega)>\delta\}$, 以及 $\varphi_\delta(x)=\delta^{-n}\varphi\left(\frac{x}{\delta}\right)$, 其中 $\varphi\in C_0^\infty(B_1(\theta))$, 满足 $\int_{B_1(\theta)}\varphi(x)dx=1$, $\varphi(x)=\varphi(-x)$. 我们有

$$u_\delta(x)=\int_\Omega u(y)\varphi_\delta(x-y)dy\in C^\infty(\Omega_\delta).$$

引理 13.2 若 $u\in C(\Omega)$ 具有平均值性质, 则 $u\in C^\infty(\Omega)$, 并且 $u(x)=u_\delta(x)$, $\forall x\in\Omega_\delta$.

证明 $\forall x\in\Omega_\delta$, 由 u 的平均值性质,

$$\begin{aligned}u_\delta(x)&=\delta^{-n}\int_{|y|\leqslant\delta}u(x+y)\varphi\left(\frac{y}{\delta}\right)dy\\&=\int_{|y|\leqslant 1}u(x+\delta y)\varphi(y)dy\\&=\int_0^1 r^{n-1}\rho(r)\int_{|w|=1}u(x+\delta rw)dwdr=u(x),\end{aligned}$$

其中 $\rho(r)=\varphi(rw),|w|=1$. 因为 $u_\delta\in C^\infty(\Omega_\delta)$ 而且 $\delta>0$ 是任意的, 所以 $u\in C^\infty(\Omega)$. □

引理 13.3 若 $u\in L^1(\Omega)$ 是一个弱调和函数, 即满足方程 (13.2), 则 $u\in C^\infty(\Omega)$, 并具有平均值性质.

证明 (1) 先证明 $\Delta u_\delta(x)=0$, $\forall x\in\Omega_{\delta_0}$, $0<\delta<\delta_0$.

$\forall\varphi\in C_0^\infty(\Omega_{\delta_0})$, 因为 $\mathrm{supp}\varphi_\delta\subset\Omega$, 所以有

$$\begin{aligned}\int_{\Omega_{\delta_0}}u_\delta(x)\Delta\varphi(x)dx&=\int_{\mathbb{R}^n}u_\delta(x)\Delta\varphi(x)dx\\&=\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\tilde{u}_\delta(y)j_\delta(x-y)\Delta\varphi(x)dx\\&=\int_\Omega u(y)(\Delta\varphi)_\delta(y)dy\\&=\int_\Omega u\Delta\varphi_\delta dy=0.\end{aligned}$$

于是, $\Delta u_\delta(x)=0$ 于 Ω_{δ_0}.

(2) 如能证明 $\{u_\delta\}$ 在 $\bar{\Omega}_{2\delta_0}$ 上是一致有界、等度连续的, 那么便存在子列一致收敛到一个连续函数 v. 因为 $u_\delta\to u$ 于 $L^1(\Omega_{\delta_0})$, 所以 $u(x)=v(x)$ a.e., 又因为 u_δ 是调和的, 所以有平均值性质, 取极限后, v 也具有平均值性质. 应用引理 13.2, 即得 $v\in C^\infty(\Omega)$.

(3) 现在回过来证明 $\{u_\delta\}$ 在 $\bar{\Omega}_{2\delta_0}$ 上是一致有界、等度连续的. 用 u_δ 的平均值性质,

$$|u_\delta(x)| \leqslant \frac{n}{\delta_0^n \varpi_n} \int_{B_{\delta_0}(x)} |u_\delta(y)| dy \leqslant C_{\delta_0} \|u_\delta\|_1 \leqslant C_{\delta_0} \|u\|_1,$$

以及

$$|u_\delta(x) - u_\delta(y)| \leqslant \left| \int_\Omega u(\xi) \int_{B_{\delta_0}(\theta)} [\varphi(x + \delta(z - \xi)) - \varphi(y + \delta(z - \xi))] d\xi dz \right|$$
$$\leqslant C_{\delta_0} |x - y| \|u\|_1.$$

□

注 13.1 定理 13.2 说明广义解在 Ω 内是光滑的, 进一步要研究它在 $\overline{\Omega}$ 上的光滑性, 后者要比前者困难得多, 已超出本讲义的范围. 研究偏微分方程解的正则性是一门专门的学问. 在这方面已有许多教材、专著和重要文献, 入门的教材可参看 D. Gilbarg, N. Trudinger 的 [GT]. 在这里我们不再展开讨论. □

对于线性强椭圆型方程, 在区域、系数都足够光滑的情况下, 正则边值问题的解有足够的可微性, 其代表性成果属于 J. Schauder, S. Agmon, A. Douglise, L. Nirenberg. 对于拟线性椭圆型方程, 则属于 J. Nash, E. De Giorgi, O. A. Ladyzenskaya, N. N. Uraltzeva. 至于椭圆型方程组的情形, 一般来说, 解未必有正则性. 例如见 M. Giaquinta 的 [Gi],

§13.3 几个变分问题的求解

我们通过几个例子看变分问题解的存在性与正则性.

例 13.4 (两点边值问题) 记区间 $J = (t_0, t_1)$, 环面 $T^2 = \mathbb{R}^2/\mathbb{Z}^2$. 设给定 $a_0, a_1 \in T^2$, $F \in C^2(J \times T^2)$. 求解 $u \in C^2(J, T^2)$, 满足边值条件 $u(t_i) = a_i, i = 0, 1$, 以及下列方程:

$$\ddot{u}(t) = \nabla_u F(t, u(t)). \tag{13.4}$$

解 为了转化到标准的变分问题, 先把环面上的函数 F 扩张到 $\mathbb{R}^2$ 上去, 即设

$$F(t, u_1 + 1, u_2) = F(t, u_1, u_2 + 1) = F(t, u_1, u_2), \quad \forall t \in J,$$

仍用 F 表示. 定义泛函及相应的空间如下:

$$I(u) = \int_J \left[\frac{1}{2}|\dot{u}(t)|^2 + F(t, u(t))\right] dt,$$

$$M = u_0 + H_0^1(J, \mathbb{R}^2), \quad u_0 = \frac{a_0(t_1 - t) + a_1(t - t_1)}{t_1 - t_0}.$$

它的 E-L 方程就是 (13.4). □

记 $u = u_0 + v$, 则模 $\|u\|$ 与 $\|u_0\| + \left(\int_J |\dot{v}|^2\right)^{1/2}$ 等价.

因为 F 有界, 所以 I 是强制的. 又因为 $H^1(J, \mathbb{R}^2) \to C(J, \mathbb{R}^2)$ 的嵌入是紧的, 而且 F 是连续的, 所以 I 是序列弱下半连续的. 此外, M 也是弱闭的. 按存在性定理, I 在 M 上有极小点 $u \in M$. 如今 $(L_{p_i p_j}) = Id$, 当然是正定的, 再按正则性定理 13.1, 这个 $u \in C^2$. 显然 u 满足 (13.4).

例 13.5 (强迫振动的周期解) 设 $e \in C[0, T]$ 是一个平均值为 0 的 T 周期函数: $\int_0^T e(t)dt = 0$, a 是一个常数. 求下列方程周期为 $T > 0$ 的解:

$$\ddot{u}(t) + a \sin u(t) = e(t).$$

解 定义 $H^1_{\text{per}}(0, T)$ 为周期为 $T > 0$ 的 Sobolev 空间 $H^1(0, T)$, 即周期为 T 的 C^∞ 函数在 $H^1(0, T)$ 下的闭包. 定义泛函

$$I(u) = \int_0^T \left(\frac{1}{2}|u'(t)|^2 + a \cos u(t) - E(t)u'(t)\right) dt,$$

其中

$$E(t) = \int_0^t e(s)ds.$$

E 也是周期为 $T > 0$ 的函数. 因为

$$\int_0^T E(t)u'(t)dt = -\int_0^T e(t)u(t)dt,$$

所以 I 的 E-L 方程就是

$$\ddot{u}(t) + a \sin u(t) = e(t).$$

为了验证 I 在 $H^1_{\text{per}}(0, T)$ 有极小值, 类似于 Poincaré 不等式, 我们需要下列引理.

引理 13.4 (Wirtinger 不等式) 设 $u \in H^1_{\text{per}}(0,T)$ 且 $\bar{u} = \frac{1}{T}\int_0^T u(t)dt = 0$, 则

$$\int_0^T |u'(t)|^2 dt \geqslant \frac{4\pi^2}{T^2}\int_0^T |u(t)|^2 dt.$$

证明 在 $[0,T]$ 上对周期函数 u 用 Fourier 级数展开. 因为 $\bar{u}=0$, 所以

$$u(t) = \sum_{k=1}^{\infty}\left(a_k \cos\frac{2\pi kt}{T} + b_k \sin\frac{2\pi kt}{T}\right),$$

于是

$$u'(t) = \frac{2\pi}{T}\sum_{k=1}^{\infty}\left(-ka_k \sin\frac{2\pi kt}{T} + kb_k \cos\frac{2\pi kt}{T}\right).$$

由 Parseval 等式

$$\begin{aligned}\int_0^T |u'(t)|^2 dt &= \frac{(2\pi)^2}{T}\sum_{k=1}^{\infty} k^2(|a_k|^2 + |b_k|^2) \\ &\geqslant \frac{(2\pi)^2}{T}\sum_{k=1}^{\infty}(|a_k|^2 + |b_k|^2 \\ &= \frac{(2\pi)^2}{T^2}\int_0^T |u(t)|^2 dt.\end{aligned}$$

□

$\forall\ u \in H^1_{\text{per}}(0,T)$, 作分解

$$u = \tilde{u} + \bar{u},$$

其中 $\bar{u} = \frac{1}{T}\int_0^T u(t)dt$ 是一个实数. 从 Wirtinger 不等式可见, $\bar{u}$ 不为泛函 I 的值所控制. 换句话说, 如果直接在空间 $H^1_{\text{per}}(0,T)$ 上用 H^1 模的话, 那么 I 不是强制的.

注意到泛函 I 中的非线性项 $\cos u$ 是 2π 周期的, 所以有

$$I(u+2\pi) = I(u).$$

这表明, 实际上不必在整个 $H^1_{\text{per}}(0,T)$ 上来考虑 I, 而是取集合

$$M = \{u = \xi + \eta \,|\, \xi \in H^1_{\text{per}}(0,T),\ \bar{\xi} = 0, \eta \in [0, 2\pi]\},$$

并把 I 限制在 M 上, 这样的好处是 $\bar{u}$ 只在有界区间 $[0,2\pi]$ 上变化.

如今, 集合 M 是序列弱闭的. 注意到

$$I(u) \geqslant \frac{1}{2}\|\dot{\xi}\|_2^2 - \|E\|_\infty \|\dot{\xi}\|_2 - |a|,$$

由 Wirtinger 不等式, $\|\dot{\xi}\|_2$ 是 M 上 H^1 模的等价模. I 在 M 上是强制的. 不难验证它还是弱下半连续的. 按存在性定理, 存在极小点 $u \in M \subset H^1_{\text{per}}$. 正则性定理 13.1 的条件也都满足, 所以 $u \in C^2$. □

例 13.6 设 $\Omega \subset \mathbb{R}^n$ 是有界区域, $1 < r < p < \infty$, $f \in L^{\frac{p}{p-r}}(\Omega)$,

$$I(u) = \int_\Omega \left(\frac{1}{p}|\nabla u(x)|^p - \frac{1}{r}f(x) \cdot |u(x)|^{r-1}u(x)\right) dx, \quad \forall u \in H^1_0(\Omega)$$

的极小值是存在的, 从而满足

$$\int_\Omega [\nabla u(x) \nabla \varphi(x) - f(x)\varphi(x)]dx = 0, \ \forall \varphi \in C_0^\infty(\Omega).$$

这是方程

$$-\Delta_p u = f|u|^{r-1}$$

的广义解, 其中

$$-\Delta_p u = \sum_{i=1}^n \partial_i(|\nabla u|^{p-2})\partial_i u$$

是 p-Laplace 算子.

证明 用记号 $\|u\| = \left(\int_\Omega |\nabla u(x)|^p dx\right)^{1/p}$, 我们知道它是 $W_0^{1,p}(\Omega)$ 的一个等价模.

(1) $I(u)$ 是强制的.

由 Poincaré 不等式、Hölder 不等式及 Young 不等式, $\exists C > 0$, 使得

$$\int_\Omega |f||u|^r dx \leqslant \|f\|_{\frac{p}{p-r}} \|u\|_p^r \leqslant C\|f\|_{\frac{p}{p-r}} \|u\|^r \leqslant C_\varepsilon \|f\|_{\frac{p}{p-r}}^{\frac{p}{p-r}} + \varepsilon\|u\|^p,$$

所以有

$I(u) \geqslant (\frac{1}{p} - \varepsilon)\|u\|^p - \frac{C_\varepsilon}{r}\|f\|_{\frac{p}{p-r}}^{\frac{p}{p-r}} \to +\infty$, 当 $\|u\| \to \infty$ 时.

(2) 因为 $\int_\Omega |\nabla u(x)|^p dx$是凸的, 而且是下半连续的, 从而是序列弱下半连续的, 后面那项按 Rellich-Kondrachov 定理, 也是序列弱下半连续的, 所以 I 是序列弱下半连续的. □

注 13.2 一般说来, 不能得到 $u \in C^2$. 但当 $p = 2, r = 1$, $f \in C^\gamma(\bar{\Omega})$ $(0 < \gamma < 1)$, 是 Hölder 函数类, 并且 $\partial\Omega$ 足够光滑时, 利用 Schauder 估计可以证明 $u \in C^{2,\gamma}(\bar{\Omega})$. 当 $p = 2, r = 1$, $f \in C(\bar{\Omega})$ 时就是 Poisson 方程. □

例 13.7 (调和映射) 设 $\Omega \subset \mathbb{R}^m$ 是一有界开区域, 有光滑的边界 $\partial\Omega$. 记 $\mathbb{R}^{n+1}$ 中的单位球为 S^n. 考虑映射 $u = (u^1, \cdots, u^{n+1}):\ \Omega \longrightarrow S^n \subset \mathbb{R}^{n+1}$. 并给定边界上的一个 C^1 映射 $\varphi = (\varphi^1, \cdots, \varphi^{n+1}):\ \partial\Omega \longrightarrow S^n$. 定义集合

$$M = \{u \in H^1(\Omega, \mathbb{R}^{n+1}) \mid u(x) \in S^n,\ \text{a.e. } x \in \Omega,\ u|_{\partial\Omega} = \varphi\},$$

求

$$\inf\{E(u) \mid u \in M\},$$

其中

$$E(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 dx.$$

因为 E 是凸的, 又是下半连续的, 所以是序列弱下半连续的. 此外 E 还是强制的.

现在我们来验证 M 是序列弱闭集. 事实上, 若

$$u_j \rightharpoonup u\ \ (H^1(\Omega, \mathbb{R}^{n+1})),$$

则存在子列 $u_{j'} \to u$, $L^2(\Omega, \mathbb{R}^{n+1})$, 又存在子列 $u_{j'}(x) \to u(x)$ a.e..

由 $u_j(x) \in S$ a.e., 推得 $u(x) \in S^n$ a.e.. 从而 $u \in M$. 于是存在极小 $u_0 \in M$ 满足下列 E-L 方程:

$$\int_\Omega \nabla u_0(x) \nabla v(x) dx = 0,$$

$\forall v \in H_0^1(\Omega, \mathbb{R}^{n+1})$, 满足 $v(x) \in T_{u_0(x)}S^n$, a.e. 于 Ω, 其中 T_uS^n 是在 $u \in S^n$ 处 S^n 的切空间.

如果 $u_0 \in C^2(\Omega, \mathbb{R}^{n+1})$, 那么

$$(\Delta u_0)^T(x) = 0, \quad \text{a.e.于 } \Omega,$$

其中 $(\Delta u_0)^T(x)$ 是 $\Delta u_0(x)$ 在 $u_0(x)$ 处的切向投影.

因为 $\Delta u_0(x)$ 在 $u_0(x)$ 处的法向投影

$$(\Delta u_0)^N(x) = \Delta u_0(x) \cdot u_0(x),$$

对 $|u(x)|^2 = 1$ 求导两次得到

$$\Delta u(x) \cdot u(x) = -|\nabla u(x)|^2,$$

所以 u_0 满足方程

$$-\Delta u(x) = u(x)|\nabla u(x)|^2.$$

这就是所谓的调和映射方程 (参看第七讲例 7.5). □

注 13.3 当 $m=2$ 时, Morrey 证明了 u_0 是光滑的. 但当 $m>2$ 时, 林芳华证明了 $u_0=1/|x|$. 可见即使解 u_0 是极小点, 它也有奇点. □

例 13.8 (非线性特征值问题) 在第十二讲我们介绍了如何用变分方法讨论线性微分方程的特征值问题, 非线性微分方程也有特征值问题, 同样可以用变分方法来研究. 举一个例子, 给定有界区域 $\Omega\subset\mathbb{R}^n$, 给定一个满足下列增长条件的 Carathéodory 函数 $f:\Omega\times\mathbb{R}^1\to\mathbb{R}^1$:

$$|f(x,u)|\leqslant C(1+|u|^r),\quad 1\leqslant r<2^*-1,$$

其中 $2^*=\frac{2n}{n-2}$ $(n>2)$. 假设

$$f(x,0)=0.$$

求 $u\in H_0^1(\Omega)\backslash\{0\}$, 满足方程

$$-\Delta u(x)=\lambda f(x,u(x)),$$

其中 λ 是一个参数. 对应有非零解 $u\in H_0^1(\Omega)$ 的 λ 就称为一个特征值.

解 和线性方程一样, 我们把它看成是一个约束变分问题. 在 $H_0^1(\Omega)$ 引入泛函

$$I(u)=\frac{1}{2}\int_\Omega|\nabla u(x)|^2dx,$$

$$N(u)=\int_\Omega F(x,u(x))dx,$$

其中

$$F(x,t)=\int_0^t f(x,s)ds$$

是 $f(\cdot,t)$ 的一个原函数. 由于对 f 做了增长性限制, 所以 $N(u)$ 在 $H_0^1(\Omega)$ 有定义.

同样我们考虑 $M=N^{-1}(1)$, 并在 M 上求 I 的极小值. 当然 I 是强制的、弱下半连续的, 再利用 Rellich-Kondrachov 紧嵌入定理, M 是弱闭的. 因此 I 存在极小点 u_0. 又按 Lagrange 乘子定理, 存在实数 λ_0 满足

$$-\Delta u_0(x)=\lambda_0 f(x,u_0(x)).$$

当 $f(x,u)=u$ 时, 这就是第十二讲讨论过的线性特征值问题. 当 $f(x,u)=u^r$, $1\leqslant r<2^*-1$, $2^*=\frac{n+2}{n-2}$ $(n>2)$, 它显然是线性特征值的推广. □

注 13.4 然而在第十二讲, 我们对于线性问题得到了一系列特征值. 对于怎样的非线性特征值问题也有类似的结论? 这种情形不像线性问题那样简单, 因为

没有正交性可以利用. 如果函数 f 对于 u 是奇的, 即 $f(x,-u)=-f(x,u)$, 类似的结果有 Liusternik-Schnirelmann 理论, 需要用到更深刻的拓扑学的知识才能得到. $\square$

我们再来看一个例子. 所研究的泛函表面上看既无上界又无下界, 似乎求极值的方法用不上, 然而通过一种特定的技术, 可以化归为求极小值的问题.

例 13.9 (极化方法) 求 $u\in X:=H_0^1(\Omega)$ 满足下列方程:

$$-\Delta u=|u|^{p-2}u,$$

其中 $2<p<2^*$.

它是下列泛函的 E-L 方程:

$$I(u)=\int_\Omega\left(\frac{1}{2}|\nabla u|^2-\frac{1}{p}|u|^p\right)dx.$$

但这泛函既无上界又无下界.

我们采用极化分解的办法把自变量分解. 设 S 为 X 的单位球面, $\forall u\in X\backslash\{\theta\}$, 存在唯一的 $(t,v)\in\mathbb{R}_+^1\times S$, 使得 $u=tv$.

任意固定 $v\in S$, 考虑定义在半直线上的一元函数

$$t\to I(tv)=\frac{t^2}{2}-\frac{t^p}{p}|v|_p^p,$$

其中 $|v|_p^P=\int_\Omega|v(x)|^Pdx$. 它有极大点

$$t=t(v)=\frac{1}{|v|_p^{\frac{p}{p-2}}},$$

满足

$$\frac{d}{dt}I(tv)|_{t=t(v)}=0.$$

代入原泛函得到一个在单位球面 S 上的泛函

$$\tilde{I}(v)=I(t(v)v)=\left(\frac{1}{2}-\frac{1}{p}\right)\frac{1}{|v|_p^{\frac{2p}{p-2}}}.$$

利用嵌入定理, 存在常数 $C>0$, 使得

$$|v|_p^p\leqslant C^p\|v\|^p=C^p.$$

从而

$$\frac{1}{|v|_p^p} \geqslant \frac{1}{C^p}.$$

这表明 $\tilde{I}$ 是 S 上定义的一个连续泛函. 再由紧嵌入，它还是弱序列连续的. 注意到 S 是弱列紧的, 所以 $\tilde{I}$ 存在极小点 v_0.

我们说 $u_0 = t(v_0)v_0$ 就是原方程的解. 事实上, 一方面我们有

$$0 = (\tilde{I}'(v_0), w) = (I'(t(v_0)v_0), w) - (I'(t(v_0)v_0), v_0)(v_0, w), \quad \forall\, w \bot v_0,$$

其中 $(\cdot,\cdot)$ 是 X 上的内积. 另一方面,

$$\frac{d}{dt} I(tv)|_{t=t(v)} = (I'(t(v)v), v).$$

所以

$$(I'(t(v_0)v_0), w + tv_0) = 0, \ \ \forall\, w \bot v_0, \forall\, t \in \mathbb{R}^1,$$

此即

$$(I'(u_0), \varphi) = 0, \quad \forall\, \varphi \in X.$$

我们证明了 u_0 是泛函 I 的 E-L 方程的解. □

注 13.5 在文献中有人把极化方法称为纤维化方法. □

§13.4 变分学的局限

在本讲的最后, 我们特别指出: 一个微分方程的解并不总是能用直接法求得的. Hadamard 曾举过下列反例. 令

$$g(\vartheta) = \sum_{m=1}^{\infty} \frac{\sin m!\vartheta}{m^2},$$

则

$$u(r, \vartheta) = \sum_{m=1}^{\infty} \frac{r^{m!} \sin m!\vartheta}{m^2} \quad ((r, \vartheta) \in [0, 1] \times [0, 2\pi])$$

在单位圆上一致收敛到一个连续函数 u.

这个函数 u 在单位圆内光滑并且是调和的, 取边值 g. 但这函数梯度平方的积分是 ∞! 换句话说, 作为调和函数 u 在某种意义下满足 Dirichlet 积分的 E-L 方程 $\Delta u = 0$, 但它却使对应的泛函 (Dirichlet 积分) 取值无穷!

习 题

1. 设 $\Omega \subset \mathbb{R}^n$ 是一个边界光滑的有界区域, 又设 $f \in C(\bar{\Omega})$. 在 $W_0^{1,4}(\Omega)$ 上考察泛函

$$I(u) = \int_\Omega \left[\frac{1}{4}|\nabla u|^4 - x^2|u|^2 - f(x)u\right] dx,$$

试证它是弱下半连续的, 是强制的, 有极小解.

2. 下列泛函是弱下半连续的吗? 是强制的吗? 存在极小解吗? 并说明理由.

(1)

$$\int_{-1}^{1} t^2\dot{u}^2 dt, \quad M = \{u \in H^1(-1,1) \,|\, u(\pm 1) = \pm 1\}.$$

(2)

$$\int_0^1 (\dot{u}^2 - 1)^2 dt, \quad M = W_0^{1,4}(0,1).$$

第十四讲　对偶作用原理与 Ekeland 变分原理

这一讲包括两方面的内容：对偶作用原理与 Ekeland 变分原理.

对偶作用原理主要应用于与 Hamilton 系统有关的问题. 因为 Hamilton 系统对应的泛函是不定的, 既无上界又无下界, 所以求极值的方法无法施展. 然而如果一个 Hamilton 系统中的 Hamilton 函数具有凸性, 那么可以通过凸函数的共轭函数把问题转化. 这就是对偶作用原理.

Ekeland 变分原理则是一个很一般、应用很普遍的极小化原理, 它提供了一种选取极小化序列的方法. 这串极小化序列结合其他条件有广泛的应用.

§14.1　凸函数的共轭函数

在第三讲中我们曾介绍过 $\mathbb{R}^n$ 上一个函数 $f(x)$ 的 Legendre 变换 f^*,

$$f^*(\xi) = \xi \cdot \psi(\xi) - f(\psi(\xi)), \tag{14.1}$$

其中 $x = \psi(\xi)$ 是 $\xi = \nabla f(x)$ 的反函数.

Lagendre 变换的重要性体现在: f 的梯度 ∇f 与它的 Legendre 变换 f^* 的梯度 ∇f^* 互为逆变换, 即

$$\xi = \nabla f(x), \quad x = \nabla f^*(\xi)$$

成互逆的关系, 因此从 f^* 我们能回到 f. 它们之间的联系是

$$f(x)+f^*(\xi)=\langle\xi,x\rangle.$$

但是, Legendre 变换可应用的范围非常有限, 因为在它的定义中要求 $(\nabla f(x))^{-1}$ 是处处存在的.

给定一个凸函数 f, 它的定义域是 $D(f)=\{x\in\mathbb{R}^n\mid f(x)<+\infty\}$. 取代梯度映射概念的是*次微分* $x\mapsto\partial f(x)$. 次微分是一个集值映射:

$$\begin{aligned}\xi\in\partial f(x)&\Longleftrightarrow f(y)\geqslant f(x)+\langle\xi,y-x\rangle,\quad\forall y\in D(f)\\&\Longleftrightarrow f(x)-\langle\xi,\,x\rangle=\min_{y\in D(f)}\{f(y)-\langle\xi,\,y\rangle\}.\end{aligned}$$

它的逆映射也是一个集值映射. 若仍以 $\psi(\xi)$ 表示 $\partial f(x)$ 的逆 (集值) 映射, 则

$$\psi(\xi)=\{x\in D(f)\,|\,\xi\in\partial f(x)\}.$$

即,

$$x\in\psi(\xi)\Longleftrightarrow x\text{ 使函数 }f(y)-\langle\xi,\,y\rangle\text{ 达到极小值}$$

这样一来, 我们就有可能把 Legendre 函数的概念扩充到凸函数.

我们把它与 (14.1) 进行比较, 引入

定义 14.1 (共轭函数) *设 $f:\mathbb{R}^n\to\bar{\mathbb{R}}=\mathbb{R}^1\cup\{+\infty\}$ 是真 (proper) 函数, 即 $D(f)\neq\varnothing$. 我们把*

$$f^*(\xi)=\sup_{x\in\mathbb{R}^n}\{\langle\xi,x\rangle-f(x)\}$$

称为 f 的共轭函数. 也有人把它称为 f 的 Fenchel 变换.

共轭函数有下列简单性质.

(1) f^* 是下半连续的函数.

这是因为 f^* 是一族仿射函数之上确界, 而仿射函数既是凸的, 又是下半连续的, 而凸下半连续的函数族的上确界还是凸下半连续的.

(2) 若 f 是真的、下半连续的凸函数, 则 f^* 是真函数.

证明 因为 f 是真的, $\exists x_0\in\mathbb{R}^n$ 使得 $f(x_0)<+\infty$. 这表明上方图 (epigraph)$\mathrm{epi}(f)=\{(x,t)\in\mathbb{R}^n\times\mathbb{R}^1\mid f(x)\leqslant t\}$ 是余集非空的闭凸集. 取 $t_0<f(x_0)$, 则 $(x_0,t_0)\notin\mathrm{epi}(f)$. 应用 Ascoli 分离定理, $\exists(x_0^*,\lambda)\in\mathbb{R}^n\times\mathbb{R}^1$, $\exists\alpha\in\mathbb{R}^1$ 使得

$$\langle x_0^*,x\rangle+\lambda t>\alpha>\langle x_0^*,x_0\rangle+\lambda t_0,\quad\forall(x,t)\in\mathrm{epi}(f).\tag{14.2}$$

特别地,

$$\langle x_0^*, x_0\rangle + \lambda f(x_0) > \alpha > \langle x_0^*, x_0\rangle + \lambda t_0.$$

从而 $\lambda > 0$, 并且

$$\left\langle -\frac{1}{\lambda}x_0^*, x\right\rangle - f(x) < -\frac{\alpha}{\lambda}, \quad \forall x \in D(f),$$

即

$$f^*\left(-\frac{x_0^*}{\lambda}\right) < -\frac{\alpha}{\lambda} < +\infty. \qquad \square$$

(3) 若 $f \leqslant g$, 则 $g^* \leqslant f^*$.

(4) (Young 不等式)$\langle x^*, x\rangle \leqslant f(x) + f^*(x^*)$.

(5) $f(x) + f^*(x^*) = \langle x^*, x\rangle \Longleftrightarrow x^* \in \partial f(x)$.

证明 由定义,

$$\begin{aligned} x^* \in \partial f(x) &\Longleftrightarrow \langle x^*, y - x\rangle \leqslant f(y) - f(x), \quad \forall y \in \mathbb{R}^n, \\ &\Longleftrightarrow \langle x^*, y\rangle - f(y) \leqslant \langle x^*, x\rangle - f(x), \quad \forall y \in \mathbb{R}^n, \\ &\Longleftrightarrow f^*(x^*) \leqslant \langle x^*, x\rangle - f(x). \end{aligned}$$

连同 Young 不等式即得结论. $\square$

(6) 若 $g(x) = f(x - x_0) + \langle x_0^*, x\rangle + a$, 则

$$g^*(x^*) = f^*(x^* - x_0^*) + \langle x^*, x_0\rangle - (a + \langle x_0^*, x_0\rangle).$$

(7) 若 $g(x) = f(\lambda x)$, $\lambda > 0$, 则 $g^*(x^*) = f\left(x^*/\lambda\right)$.

只要 f^* 是真函数, 我们就可以再定义它的共轭函数 f^{**}. 由性质 (1), f^{**} 还是真的、凸的、下半连续的函数.

定理 14.1 (Fenchel-Moreau 定理) *若 f 是一个真的、下半连续的凸函数, 则 $f^{**} = f$.*

证明 由 Young 不等式, 有 $f^{**}(x) \leqslant f(x)$, $\forall x \in \mathbb{R}^n$.

现在我们来证明反过来的不等式. 倘若不然, $\exists x_0 \in \mathbb{R}^n$ 使得 $f^{**}(x_0) < f(x_0)$. 类似于性质 (2) 的证明得到点 $(x_0, f^{**}(x_0)) \notin \mathrm{epi}(f)$, 从而有不等式 (14.2), 其中 $t_0 = f^{**}(x_0)$.

如果 $f(x_0) < +\infty$, 那么 (14.2) 中的 $\lambda > 0$, 并且

$$\langle x_0^*, x\rangle + \lambda f(x) > \alpha, \quad \forall x \in D(f).$$

我们同样得到

$$f^*\left(-\frac{x_0^*}{\lambda}\right) \leqslant -\frac{\alpha}{\lambda}.$$

按 f^{**} 的定义

$$f^{**}(x_0) \geqslant \left\langle -\frac{x_0^*}{\lambda}, x_0\right\rangle - f^*\left(-\frac{x_0^*}{\lambda}\right),$$

从而

$$\langle x_0^*, x_0\rangle + \lambda f^{**}(x_0) \geqslant \alpha. \tag{14.3}$$

这与 (14.2) 矛盾.

如果 $f^{**}(x_0) < +\infty$, 同样蕴含了 $\lambda > 0$, 那么依旧导出矛盾.

剩下 $f(x_0) = f^{**}(x_0) = +\infty$ 以及 $\lambda = 0$ 的情形. 由 (14.2), $\exists \varepsilon > 0$ 使得

$$\langle x_0^*, x - x_0\rangle \geqslant \varepsilon, \quad \forall x \in D(f). \tag{14.4}$$

因为 f^* 是真的, 所以 $\exists x_1^* \in \mathbb{R}^n$, 使得 $f^*(x_1^*) < +\infty$, 而且

$$\langle x_1^*, x\rangle - f(x) - f^*(x_1^*) \leqslant 0, \quad \forall x \in D(f). \tag{14.5}$$

联合 (14.4) 与 (14.5), $\forall n \in \mathbb{N}$, 有

$$\langle x_1^* - nx_0^*, x\rangle + n\langle x_0^*, x_0\rangle + n\varepsilon - f(x) - f^*(x_1^*) \leqslant 0, \quad \forall x \in D(f),$$

即得

$$f^*(x_1^* - nx_0^*) + n\langle x_0^*, x_0\rangle + n\varepsilon - f^*(x_1^*) \leqslant 0,$$

或

$$n\varepsilon + \langle x_1^*, x_0\rangle - f^*(x_1^*) \leqslant \langle x_1^* - nx_0^*, x_0\rangle - f^*(x_1^* - nx_0^*) \leqslant f^{**}(x_0).$$

令 $n \to \infty$, 得到 $f^{**}(x_0) = +\infty$, 这与假设矛盾. □

推论 14.1 对于一个真的、下半连续的凸函数 f, 我们有

$$\xi \in \partial f(x) \Longleftrightarrow x \in \partial f^*(\xi).$$

这是 Fenchel-Moreau 定理与性质 (5) 的直接推广.

推论 14.2 设 $f:\mathbb{R}^n\to\bar{\mathbb{R}}$ 是一个真函数, 则

$$f^{**}=\operatorname{conv}(f)=\sup\{\varphi:\mathbb{R}^n\to\bar{\mathbb{R}}\mid\varphi(x)\leqslant f(x),\ \forall x\in\mathbb{R}^n,\ \varphi\ \text{是凸函数}\}.$$

证明 设 $g\leqslant f$ 是凸的, 则它是真、凸的. 由性质 (3), $f^*\leqslant g^*$, 并且 $g^{**}\leqslant f^{**}$. 再由 Fenchel-Moreau 定理, $g=g^{**}\leqslant f^{**}$. □

例 14.1 若 $f(x)=|x|^p/p$, $1<p<\infty$, 其中 $|x|=(x_1^2+x_2^2+\cdots+x_n^2)^{1/2}$, 则

$$f^*(\xi)=\frac{1}{p'}|\xi|^{p'},\quad \text{其中}\ \frac{1}{p}+\frac{1}{p'}=1.$$ □

例 14.2 若 $f(x)=|x|$, 则

$$f^*(\xi)=\begin{cases}0, & |\xi|\leqslant 1,\\ +\infty, & |\xi|>1.\end{cases}$$

证明 事实上,

$$f^*(\xi)=\sup_x\{\langle\xi,x\rangle-|x|\}.$$

当 $|\xi|>1$ 时, 取 $x=t\xi$, 并令 $t\to+\infty$, 则 $f^*(\xi)=+\infty$. 当 $|\xi|\leqslant 1$ 时, 因 $\langle\xi,x\rangle-|x|\leqslant 0$, 再取 $x=0$, 即得结论. □

§14.2 对偶作用原理

给定一个凸的 Hamilton 函数 $H\in C^1(\mathbb{R}^N\times\mathbb{R}^N)$ 满足下列增长条件:

$$0\leqslant H(u,\xi)\leqslant C(|u|^2+|\xi|^2). \tag{14.6}$$

求一个向量值周期函数 $(u(t),\xi(t))\in\mathbb{R}^N\times\mathbb{R}^N$ 满足 Hamilton 方程组

$$\begin{cases}\dot{\xi}(t)=H_u(u(t),\xi(t)),\\ \dot{u}(t)=-H_\xi(u(t),\xi(t)).\end{cases} \tag{14.7}$$

我们说过, 极小化方法无法直接应用于它对应的泛函. 我们要用凸函数 H 的共轭函数 H^* 把这个问题转化.

在这个问题中, 周期是待定的. 我们引进一个参数 $\lambda>0$, 来求下列方程的一个以 2π 为周期的解 $(v(t),\eta(t))\in\mathbb{R}^N\times\mathbb{R}^N$:

$$\begin{cases}\dot{\eta}(t)=\lambda H_v(v(t),\eta(t)),\\ \dot{v}(t)=-\lambda H_\eta(v(t),\eta(t)).\end{cases} \tag{14.8}$$

事实上, 如果已经求得 $(v(t),\eta(t))$ 满足 (14.8), 那么令

$$u(t)=v(\lambda^{-1}t),\quad \xi(t)=\eta(\lambda^{-1}t),$$

它们满足 (14.7). 如果 $\lambda>0$, 那么它们以 $2\lambda\pi$ 为周期.

现在我们把问题 (14.8) 转化为一个约束极小值问题, 它是原问题的一个对偶问题：求 $(w,\rho)\in X:=H^1_{per}(0,2\pi)^2=\{(w,\rho)\in H^1(0,2\pi)\,|\,w(0)=w(2\pi),\ \rho(0)=\rho(2\pi)\}$, 满足

$$I(w,\rho)=\int_0^{2\pi}H^*(\dot\rho(t),-\dot w(t))dt,$$

$$G(w,\rho)=\int_0^{2\pi}(-\dot\rho(t)\cdot w(t)+\dot w(t)\cdot\rho(t))dt=-\pi. \tag{14.9}$$

使用记号

$$\langle(w,\rho),(u,\xi)\rangle=\int_0^{2\pi}(w\cdot u+\rho\cdot\xi)dt.$$

直接计算得到

$$\langle G'((w,\rho)),(u,\xi)\rangle=2\langle(-\dot\rho,\dot w),(u,\xi)\rangle=2\langle(-\dot\xi,\dot u),(w,\rho)\rangle,$$

以及

$$\langle I'((w,\rho)),(u,\xi)\rangle=\int_0^{2\pi}(H^*)'(\dot\rho(t),-\dot w(t))\cdot(\dot\xi(t),-\dot u(t))dt.$$

我们有如下结论.

(1) $M=G^{-1}(-\pi)\neq\varnothing$.

这是因为

$$G(w,\rho)=-2\int_0^{2\pi}\dot\rho(t)\cdot w(t)dt$$

对于 w 和 ρ 都是线性的.

(2) $G'((w,\rho))\neq(\theta,\theta),\forall\,(w,\rho)\in M$.

这是因为从 (14.9) 可见在 M 上 $w\neq\theta$, $\rho\neq\theta$.

(3) I 是一个下方有界的凸泛函, 在 X 上是连续的, 从而也是序列弱下半连续的.

我们只需验证它是下方有界的. 事实上, 联合 (14.6)、共轭函数的性质 (3) 与例 14.1, 可得

$$H^*(w,\rho)\geqslant\frac{1}{C}(|w|^2+|\rho|^2),$$

以及

$$I(w,\rho) \geqslant \frac{1}{C}\int_0^{2\pi}(|\dot{w}|^2+|\dot{\rho}|^2)dt.$$

注意到在 M 上没有常值向量函数, 所以

$$\|(w,\rho)\| = \left(\int_0^{2\pi}(|\dot{w}|^2+|\dot{\rho}|^2)dt\right)^{1/2}$$

是一个等价模. 这样我们同时也验证了如下结论:

(4) I 是强制的.

(5) M 是弱序列闭的.

事实上, 若 $\{(w_j,\rho_j)\}\subset M$, 在 X 上 $(w_j,\rho_j)\rightharpoonup(w_0,\rho_0)$, 则在 $L^2([0,2\pi],\mathbb{R}^N)$ 中有强收敛: $w_j\to w_0$, $\rho_j\to\rho_0$, 并且

$$\begin{aligned}
-\pi &= \lim G(w_j,\rho_j)\\
&= \lim\int_0^{2\pi}(-\dot{\rho}_j(t)\cdot w_j(t)+\dot{w}_j(t)\cdot\rho_j(t))dt\\
&= \int_0^{2\pi}(-\dot{\rho}(t)\cdot w(t)+\dot{w}(t)\cdot\rho(t))dt = G(w_0,\rho_0).
\end{aligned}$$

于是存在 $(w_0,\rho_0)\in M$ 成为约束极值问题 (14.9) 的极小点. 再应用 Lagrange 乘子定理, 存在 $\lambda\in\mathbb{R}^1$, 满足

$$I'(w_0,\rho_0)+\frac{\lambda}{2}G'(w_0,\rho_0)=0.$$

即

$$(H^*)'(\dot{\rho}_0,-\dot{w}_0)=\lambda(w_0,\rho_0). \tag{14.10}$$

利用共轭函数的次梯度与原来函数的次梯度之间的互逆关系, 即推论 14.1, 方程 (14.10) 等价于

$$\begin{cases}\dot{\rho}_0(t)=H_w(t,\lambda w_0(t),\lambda\rho_0(t)),\\ \dot{w}_0(t)=-H_\rho(t,\lambda w_0(t),\lambda\rho_0(t)).\end{cases} \tag{14.11}$$

再令

$$\begin{cases}\eta=\lambda\rho_0,\\ v=\lambda w_0,\end{cases}$$

代入 (14.11), 即是 (14.8).

最后验证 $\lambda>0$. 在 (14.10) 的两边同乘以 $(\dot{\rho}_0,-\dot{w}_0)$, 再积分, 得到

$$\langle (H^*)'(\dot{\rho}_0,-\dot{w}_0),(\dot{\rho}_0,-\dot{w}_0)\rangle = -\lambda G(w_0,\rho_0)=\lambda\pi.$$

由 (14.6) 可见 $\nabla H(\theta,\theta)=(\theta,\theta)$. 再利用共轭函数的性质 5,

$$H^*(\theta,\theta)=-H(\theta,\theta)=0.$$

由于 H^* 是一个凸函数, 所以

$$H^*(\theta,\theta)-H^*(\dot\rho_0,-\dot w_0)\geqslant -\langle (H^*)'(\dot\rho_0,-\dot w_0),(\dot\rho_0,-\dot w_0)\rangle,$$

即

$$\langle (H^*)'(\dot\rho_0,-\dot w_0),(\dot\rho_0,-\dot w_0)\rangle\geqslant H^*(\dot\rho_0,-\dot w_0)\geqslant 0.$$

我们得到 $\lambda\geqslant 0$.

剩下来我们验证 $\lambda\neq 0$. 用反证法, 倘若其为 0, 那么必有

$$(H^*)'(\dot\rho_0,-\dot w_0)=(\theta,\theta).$$

再利用推论 14.1, 可见

$$\begin{cases}\dot\rho_0=H_u(\theta,\theta),\\ \dot w_0=-H_\rho(\theta,\theta).\end{cases}$$

注意到条件 (14.6), 推出 $(H_u(\theta,\theta),H_\rho(\theta,\theta))=(\theta,\theta)$. 这与结论 (2) 矛盾!

§14.3 Ekeland 变分原理

前面我们介绍过的求极值的方法, 除了一些特殊的情形 (主要是线性问题) 可以采用正交投影外, 都离不开弱收敛 (弱序列下半连续性、弱列紧性等). 而弱拓扑是比较复杂的, 不那么容易掌握, 验证起来也比较麻烦. 下面将要介绍的 Ekeland 变分原理, 是 Ekeland 于 1970 年提出来的一条基本定理. 看来比较简单, 然而若与近代变分学常用的紧性条件 Palais-Smale 条件结合起来, 则提供了另一个非常有用的求极值的方法.

定理 14.2 (Ekeland) 设 (X,d) 是一个完备的度量空间. 又设 $f: X\to\mathbb{R}^1\cup\{+\infty\}$ 是真函数, 即 $f\not\equiv+\infty$. 若 f 下方有界并且是下半连续的, 又若 $\exists\,\varepsilon>0, \exists\, x_\varepsilon\in X$, 使得 $f(x_\varepsilon)<\inf\limits_X f+\varepsilon$, 则 $\exists\, y_\varepsilon\in X$ 使得

(1) $f(y_\varepsilon)\leqslant f(x_\varepsilon)$,

(2) $d(y_\varepsilon,x_\varepsilon)\leqslant 1$,

(3) $f(x)>f(y_\varepsilon)-\varepsilon d(y_\varepsilon,x), \forall x\in X\backslash\{y_\varepsilon\}$.

证明 在这个定理中 y_ε 是一个依赖于自身 y_ε 的函数 $f(x)+\varepsilon d(y_\varepsilon,x)$ 的极小点.

1° 我们用递归的办法确定出一个收敛序列.

先选择 $u_0 = x_\varepsilon$. 假设 u_n 已经选定, 令

$$S_n = \{w \in X \mid f(w) \leqslant f(u_n) - \varepsilon d(w, u_n)\}.$$

因为 $u_n \in S_n$, 所以 $S_n \neq \varnothing$. 选择 $u_{n+1} \in S_n$ 满足

$$f(u_{n+1}) - \inf_{S_n} f \leqslant \frac{1}{2}\left[f(u_n) - \inf_{S_n} f\right], \quad n = 0, 1, 2, \cdots.$$

2° 如此得到的序列 $\{u_n\}$ 是一个 Cauchy 列. 事实上,

$$\varepsilon d(u_n, u_m) \leqslant f(u_n) - f(u_m), \quad \forall m \geqslant n. \tag{14.12}$$

因为 $\{f(u_n)\}$ 是递减的, 而且 f 是下方有界的, 所以当 $n, m \to \infty$ 时, $f(u_n) - f(u_m) \to 0$.

从而存在 $u^* \in X$, 使得 $u_n \to u^*$. 再由 f 的下半连续性, 我们得到

$$f(u^*) \leqslant \varliminf_{n\to\infty} f(u_n) \leqslant \varliminf_{n\to\infty} \inf_{S_n} f. \tag{14.13}$$

3° 剩下来我们验证 $y_\varepsilon = u^*$ 满足 (1)–(3).

由于 $\{f(u_n)\}$ 是递减的,

$$f(y_\varepsilon) = f(u^*) \leqslant f(u_n) \leqslant f(u_0) = f(x_\varepsilon), \quad \forall n,$$

从而 (1) 成立.

又由 (14.12),

$$\begin{aligned}\varepsilon d(x_\varepsilon, y_\varepsilon) &= \varepsilon d(u_0, u^*) \\ &\leqslant f(x_\varepsilon) - f(u^*) \\ &\leqslant f(x_\varepsilon) - \inf_X f < \varepsilon,\end{aligned}$$

结论 (2) 成立.

最后用反证法证明结论 (3), 若 $y_\varepsilon = u^*$ 不满足 (3), 则 $\exists\, w \neq u^*$ 使得

$$f(w) \leqslant f(u^*) - \varepsilon d(u^*, w). \tag{14.14}$$

由 (14.12) 可得

$$\varepsilon d(u_n, u^*) \leqslant f(u_n) - f(u^*),$$

即

$$f(u^*) \leqslant f(u_n) - \varepsilon d(u_n, u^*). \tag{14.15}$$

联合 (14.14) 和 (14.15) 推得

$$f(w) \leqslant f(u_n) - \varepsilon d(u_n, w), \quad \forall n,$$

从而, $w \in \bigcap_{n=1}^{\infty} S_n$, 再由 (14.13), $f(u^*) \leqslant f(w)$, 这与 (14.14) 矛盾. □

推论 14.3 设 (X, d) 是一个完备度量空间, 又设 $f : X \to \mathbb{R}^1 \cup \{+\infty\}$ 是真的、下方有界的、下半连续的函数, 则 $\forall \varepsilon > 0$, $\exists y_\varepsilon \in X$, 使得 $f(x) > f(y_\varepsilon) - \varepsilon d(x, y_\varepsilon), \forall x \neq y_\varepsilon$.

在 Ekeland 变分原理中, 虽然只用到度量拓扑, 不涉及弱拓扑, 更不涉及各种紧性, 但是并没有得到泛函 f 的极小值. 它的意义在于, 从这串特殊选择的极小化序列提供了一串特殊的近似极小点.

§14.4 Fréchet 导数与 Palais-Smale 条件

下面我们要把 Ekeland 变分原理与连续可微函数的导数联系起来.

我们在第十讲引进过 Banach 空间上实值函数 Gâteaux 导数的概念. 设 $f : X \to \mathbb{R}^1$, $x_0 \in U \subset X$, U 是一个开邻域. 我们称 f 在点 x_0 处有 Gâteaux 导数, 如果 $\forall h \in X$, $\exists df(x_0, h) \in \mathbb{R}^1$, 使得

$$f(x_0 + th) - f(x_0) - tdf(x_0, h) = o(t),\ \ t \to 0,\ \ \forall x_0 + h \in U.$$

与 Gâteaux 导数密切相关的还有 Fréchet 导数. 如果

$$|f(x) - f(x_0) - \langle \xi, x - x_0 \rangle| = o(\|x - x_0\|) \quad (x \to x_0),$$

称 $\xi \in X^*$ 是 f 在 x_0 处的 Fréchet 导数. 这时记作 $f'(x_0) = \xi$.

如果 f 有 Fréchet 导数 $f'(x_0)$, 那么它必有 Gâteaux 导数,

$$df(x_0, h) = \langle f'(x_0), h \rangle, \quad \forall h \in X,$$

并且有

$$\|f'(x_0)\| = \sup_{h \in H} \frac{|df(x_0, h)|}{\|h\|}.$$

反过来, 设 f 在 x_0 的一个邻域 U 内处处有 Gâteaux 导数 $df(x, h)$, $x \in U$, 并且有 $\xi(x) \in X^*$, 满足

$$\langle \xi(x), h \rangle = df(x, h).$$

如果 $x \to \xi(x)$ 还是连续的, 那么 f 在 x_0 有 Fréchet 导数$f'(x_0)$.

如果 Fréchet 导数 $f'(x)$ 处处存在, 并且 $x \to f'(x)$ 是连续的, 那么我们说 f 是连续可微的. 记作 $f \in C^1(X, \mathbb{R}^1)$.

Gâteaux 导数与 Fréchet 导数分别是 $\mathbb{R}^n$ 中的方向导数与全微分在 Banach 空间的推广.

在第十一讲中我们计算过具体积分形式的泛函的 Gâteaux 导数, 它在表现形式上与变分导数是一样的, 至于 Fréchet 导数, 只要它存在, 当然也具同一形式.

例 14.3 给定 $\Omega \subset \mathbb{R}^n$, $F \in C^1(\Omega \times \mathbb{R}^1)$ 满足

$$|F_s(x,s)| \leqslant C(1+|x|)^\mu, \ \ \mu \leqslant 2^* - 1 = \frac{n+2}{n-2} \quad (n > 2).$$

在空间 $H_0^1(\Omega)$ 上求泛函

$$I(u) = \int_\Omega \left[\frac{1}{2}|\nabla u(x)|^2 + F(x, u(x))\right] dx$$

的 Fréchet 导数.

我们已经知道它的 Gâteaux 导数是

$$dI(u,v) = \int_\Omega [\nabla u \cdot \nabla v + F_u(x,u)v] dx.$$

用 $H_0^1(\Omega)$ 空间的内积

$$\langle u, v\rangle = \int_\Omega \nabla u \cdot \nabla v dx$$

把它表示出来:

$$dI(u,v) = \langle u + KF_u(x,u), v\rangle,$$

其中

$$K = (-\Delta)^{-1} : H_0^1(\Omega)^* \to (H_0^1(\Omega)).$$

由嵌入映射

$$H_0^1(\Omega) \hookrightarrow L^{2^*}(\Omega) \tag{14.16}$$

以及增长性条件

$$|F_s(x,u(x))| \leqslant C(1+|u(x)|)^\mu \in L^{\frac{2n}{n+2}} = (L^{2^*}(\Omega))^*,$$

可见非线性映射 $u \to F_s(x,u(x)) : H_0^1(\Omega) \hookrightarrow (L^{2^*}(\Omega))^*$ 是连续的.

再由嵌入映射 (14.16) 的连续性导出其共轭映射

$$(L^{2^*}(\Omega))^* \hookrightarrow (H_0^1(\Omega))^* \tag{14.17}$$

也是连续的, 于是 I 的 Gâteaux 导数 $u \to u + KF_u(x,u) : H_0^1(\Omega) \hookrightarrow (H_0^1(\Omega))$ 是连续的. 由此可见 I 的 Fréchet 导数存在并且就是

$$I'(u) = u + KF_u(x,u).$$

□

我们把使得 $f'(x_0) = \theta$ 的点 x_0 称为临界点, 并把相应的函数值 $f(x_0)$ 称为临界值.

因此在变分问题中, 极小点是临界点. 一切临界点都是 E-L 方程的解.

定义 14.2 设 X 是一个 Banach 空间, 又设 $f \in C^1(X, \mathbb{R}^1)$. 如果任意满足

$$f(x_j) \to c, \quad \|f'(x_j)\| \to 0 \tag{14.18}$$

的序列 $\{x_j\}_1^\infty \subset X$ 都有收敛的子列, 那么我们说 f 在值 c 满足 Palais-Smale 条件, 记作 PS_c. 我们还把满足 (14.18) 的序列称为一个 Palais-Smale 序列 (简称 PS 序列).

如果 $\forall c \in \mathbb{R}^1$, f 都满足 PS_c, 那么我们说 f 满足 PS.

Palais-Smale 条件的概念可以同样扩充到一般的 Banach 流形上去.

下面的推论极为重要, 它结合 Ekeland 变分原理与 Palais-Smale 条件给出存在极小值的另一个判据.

推论 14.4 设 X 是一个 Banach 空间 (或更一般的 Banach 流形), $f \in C^1(X, \mathbb{R}^1)$, 并且是下方有界的. 记

$$c = \inf_X f.$$

如果 f 还满足 PS_c, 那么 f 能达到极小值.

证明 按 Ekeland 变分原理, $\forall n \geqslant 1, \exists x_n \in X$ 满足

$$\begin{cases} f(x) > f(x_n) - \dfrac{1}{n}\|x - x_n\|, & \forall x \neq x_n, \\ f(x_n) < c + \dfrac{1}{n}. \end{cases}$$

第一个不等式意味着

$$\|f'(x_n)\| = \sup_{\|\varphi\|=1} |df(x_n, \varphi)| \leqslant \frac{1}{n}.$$

第二个不等式意味着

$$f(x_n) \to c.$$

按 PS_c, $\{x_n\}$ 有收敛子列 $\{x_{n_j}\}$: $x_{n_j} \to x^* \in X$, 再由连续性, $f(x^*) = c$. □

§14.5 Nehari 技巧

有许多泛函既无上界又无下界, 表面上看求极值的方法不能用来寻求临界点. 但对于一些特殊的问题, Nehari 使用了一种特殊的技巧, 把求临界点转化为求极值问题.

设 H 是一个 Hilbert 空间, 有内积 $(\cdot,\cdot)$, 给定一个泛函 $I\in C^2(H,\mathbb{R}^1)$. 我们要求 I 的临界点, 即 $I'(u)=0$ 的点.

定义 $G(u)=(I'(u),u)$. 注意到 I 所有的临界点 u 都满足:

$$G(u)=0$$

如果集合 $M=\{u\in H\,|\,G(u)=0\}$ 是一个流形, 例如, $G'(u)\neq\theta, \forall\, u\in M$, 那么把 I 限制在 M 上得到一个新的泛函 $\tilde{I}$. 如果 $\tilde{I}$ 有极值的话, 那么我们就来对 $\tilde{I}$ 求极值, 也就是, 求 I 的条件极值.

人们自然会问, 约束 $G(u)=0$ 产生的 Lagrange 乘子怎么办? 事实上, 因为

$$\tilde{I}'(u)=I'(u)-\frac{(I'(u),u)}{\|G'(u)\|^2}G'(u),$$

所以在 M 上 $(I'(u),u)=G(u)=0$, 只要 $G'(u)\neq\theta$, 就有

$$\tilde{I}'(u)=0\Longleftrightarrow I'(u)=0.$$

我们用一个具体例子来说明怎样运用这个技巧.

例 14.4 设 $\Omega\subset\mathbb{R}^n$ 是一个有界区域, 设 $a\in C(\bar{\Omega})$, $a(x)\geqslant\alpha>0$, $2<\mu<2^*$. 在空间 $H_0^1(\Omega)$ 上求泛函

$$I(u)=\int_\Omega\left[\frac{1}{2}|\nabla u(x)|^2+a(x)|u(x)|^\mu\right]dx$$

的非平凡临界点.

计算

$$G(u)=\langle I'(u),u\rangle=\int_\Omega[|\nabla u|^2+\mu a(x)|u(x)|^\mu]dx.$$

在 $G^{-1}(0)$ 上,

$$\tilde{I}(u)=(\mu-2)\int_\Omega a(x)|u(x)|^\mu dx=\left(\frac{1}{2}-\frac{1}{\mu}\right)\|u\|^2$$

是非负的.

1° 注意到

$$\langle G'(u), v\rangle = \int_\Omega [2\nabla u \cdot \nabla v + \mu^2 a(x)|u(x)|^{\mu-2}u(x)v(x)]dx, \quad \forall\, v \in H_0^1(\Omega),$$

可见 $G(\theta) = 0$. 并且

$$G'(u) = 2u + (-\Delta)^{-1}(\mu^2 a|u|^{\mu-2}u),$$

在 $G^{-1}(0)$ 上条件 $G'(u) \neq \theta$ 不成立.

然而 $\forall u \in G^{-1}(0)$, 应用嵌入定理有

$$\int_\Omega |\nabla u|^2 dx = \mu \left|\int_\Omega a(x)|u(x)|^\mu dx\right| \leqslant C\left(\int_\Omega |\nabla u|^2 dx\right)^{\mu/2},$$

推出或者 $u = \theta$, 或者

$$\left(\int_\Omega |\nabla u|^2 dx\right)^{\frac{\mu}{2}-1} \geqslant \frac{1}{C}.$$

这说明 θ 在 $G^{-1}(0)$ 中是一个孤立的点. 令 $M = G^{-1}(0)\backslash\{\theta\}$, 则在 M 上,

$$\langle G'(u), u\rangle = 2\|u\|^2 + \mu^2 \int_\Omega a(x)|u(x)|^\mu dx > 0,$$

可见 $G'(u) \neq 0,\ \forall u \in M$.

2° 容易看出 $\tilde{I} \in C^1(H_0^1(\Omega), \mathbb{R}^1)$, 再来验证 Palais-Smale 条件. 设有 PS 序列 $\{u_j\} \subset M$, 满足 $\tilde{I}'(u_j) \to 0$, $|\tilde{I}(u_j)| \leqslant C$. 事实上, 由于

$$\left(\frac{1}{2} - \frac{1}{\mu}\right)\|u_j\|^2 = \tilde{I}(u_j) \leqslant C,$$

所以存在子列 $u_{j'} \rightharpoonup u$. 再由

$$\tilde{I}'(u) = I'(u) = u + (-\Delta)^{-1}\mu a(x)|u(x)|^{\mu-2}u(x)$$

以及嵌入映射

$$H_0^1(\Omega) \hookrightarrow L^\mu(\Omega)$$

的紧性, 可见当 $\tilde{I}'(u_j) \to \theta$ 时, $\{u_j\}$ 有子列在 $L^\mu(\Omega)$ 中收敛. 再复合上 $(-\Delta)^{-1}$, 即得这个子序列在 $H_0^1(\Omega)$ 中收敛到 $u_0 \in M$, 并且

$$I'(u_0) = \widetilde{I}'(u_0) = \lim \widetilde{I}(u_j) = 0.$$

所以 u_0 就是所要的非平凡临界点. □

习 题

1. 设 $R > a > 0,\ M = \{u \in C^1([-a,a]) \,|\, u(\pm a) = \sqrt{R^2 - a^2}\}$,

$$I(u) = \int_{-a}^{a} \left(\sqrt{1 - \dot{u}^2} - \frac{u}{R}\right) dt.$$

(1) 计算一次、二次变分.

(2) 写出 Euler-Lagrange 方程.

(3) 验证 $u_0 = \sqrt{R^2 - t^2}$ 是其弱极小解.

(4) 写出沿 u_0 的 Jacobi 算子.

2. 求 $\inf\{I(u) \,|\, u \in M\}$:

(1)

$$I(u) = \int_1^2 t\sqrt{1 + \dot{u}^2}dt,$$

$$M = \{u \in C^1([1,2]) | u(j) = \operatorname{arcosh} j, j = 1, 2\}.$$

(2)

$$I(u) = \pi \int_a^b u^2 dt,$$

$$M = \left\{u \in C_0^1([a,b]) \,\middle|\, 2\pi \int_a^b u\sqrt{1 + \dot{u}^2}dt = c\right\}.$$

3. 设 $V \in C^1(\mathbb{R}^1, \mathbb{R}^1), (u,p) \in \mathbb{R}^N \times \mathbb{R}^N$, 给定

$$L(u,p) = \frac{1}{2}\sum_{i=1}^{N} p_i^2 - \sum_{i \neq j} V((u_i - u_j)^2).$$

(1) 写出相应的泛函.

(2) 写出 Euler-Lagrange 方程.

(3) 写出相应的 Hamilton 量.

(4) 写出相应的 Hamilton 方程组.

(5) 设 $\{u_i(t)\}_1^N$ 是 Euler-Lagrange 方程的解, 求证:

$$\frac{1}{2}\sum_{i=1}^{N} \dot{u}_i^2(t) + \sum_{i \neq j} V((u_i(t) - u_j(t))^2) = \text{const}, \quad t \in \mathbb{R}^1.$$

(6) 写出对应的 Hamilton-Jacobi 方程.

4. 设 $r \in C([0,1]), \exists t_0 \in [0,1]$, 使得 $r(t_0) > 0$. 求证存在无穷多对 $\{(\lambda_n, u_n)\}_1^\infty$, 其中 $u_n \neq 0, \lambda_n \to +\infty$, 满足

$$\begin{cases} \ddot{u} = \lambda r u \quad 于\ (0,1), \\ u(0) = u(1) = 0. \end{cases}$$

第十五讲　山路定理及其推广与应用

在这一讲我们介绍寻求极小 (大) 点以外的其他类型临界点的理论. 这个理论的基本出发点是通过泛函水平集合拓扑结构的变化来判定临界点的存在性. 因此它的理论有赖于代数拓扑学和微分拓扑学的结论和工具. 自 20 世纪 70 年代以来临界点理论有了很大的发展, 特别是在有变分结构的偏微分方程和动力系统中取得重要的应用.

我们没有假定读者已经具备这些拓扑学的知识, 本课程也只是想向读者介绍一些最基本的、最常用的临界点定理 (如山路定理等), 所以为便于读者接受, 我们尽量利用几何直观.

通常的临界点理论采用梯度流来实现泛函水平集间的形变, 这些思想与本课程内容相距很远, 需要较多的篇幅, 超出了本课程的范围. 在这一讲我们采用一种直接的方法 —— 以 Ekeland 变分原理为出发点来介绍这些临界点定理.

§15.1　山路 (Mountain Pass) 定理

一个几何上非常直观的例子提供了研究鞍点型临界点的基本思想.

在一个四面环山的盆地，从山外地面上一点 p_1 出发要想进入盆地中的一点 p_0, 人们希望走的山路是这样一条连接 p_0 与 p_1 的道路, 其最高点不高于任何邻近道路的最高点. 这条山路上的最高点就可能是一个鞍点 —— 既非极大也非极小的临界点 (见图 15.1).

用数学语言描述如下：设空间 $X = \mathbb{R}^n, \Omega \subset X$ 是一个开集. 有两点 $p_0 \in \Omega$

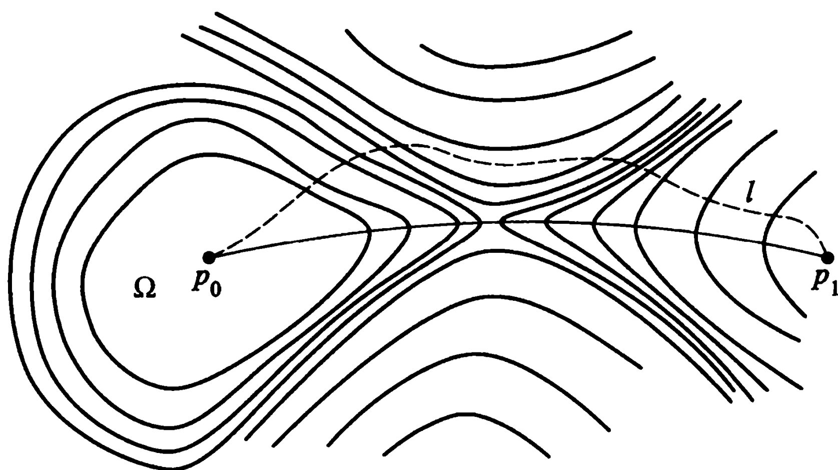


图 15.1

和 $p_1 \notin \bar{\Omega}$. 假设有一个函数 $f \in C^1(X, \mathbb{R}^1)$ 满足

$$\alpha = \inf_{x \in \partial\Omega} f(x) > \max\{f(p_0), f(p_1)\}, \tag{15.1}$$

令

$$\Gamma = \{l \in C([0,1], X) \mid l(i) = p_i, i = 0, 1\} \tag{15.2}$$

为连接这两点的所有路径组成的集合; 再令

$$c = \inf_{l \in \Gamma} \sup_{t \in [0,1]} f \circ l(t). \tag{15.3}$$

我们想说明 c 是 f 的一个临界值, 即存在 $x_0 \in X$, 使得 $f'(x_0) = 0$ 以及 $f(x_0) = c$.

然而几何直观并不等于一个确切的定理. 为了使这个结论成立, 还需要对 f 添加其他假设.

定理 15.1 (山路定理) 设 X 是一个 Banach 空间, $f \in C^1(X, \mathbb{R}^1)$; 又设 $\Omega \subset X$ 是一个开集, 给定两点 $p_0 \in \Omega$, $p_1 \notin \bar{\Omega}$, 满足 (15.1). 按 (15.2) 与 (15.3) 定义 c. 如果 f 还满足 PS_c, 那么 $c \geqslant \alpha$ 是 f 的一个临界值.

证明 在 Γ 上引入度量

$$d(l_1, l_2) = \max_{t \in [0,1]} \|l_1(t) - l_2(t)\|,$$

使 (Γ, d) 成为一个完备度量空间. 令

$$I(l) = \max_{t \in [0,1]} f \circ l(t).$$

由假设 (15.1), $I(l) \geqslant \alpha$, 并且 I 满足局部 Lipschitz 条件:

$$\begin{aligned}|I(l_1)-I(l_2)| &\leqslant \max_{t\in[0,1]}|f\circ l_1(t)-f\circ l_2(t)|\\ &\leqslant \max_{t,\theta\in[0,1]}\|f'(\theta l_1(t)+(1-\theta)l_2(t))\|\|l_1(t)-l_2(t)\|\\ &\leqslant Cd(l_1,l_2),\end{aligned}$$

其中 C 是依赖于 f', l_1, l_2 的常数. 应用 Ekeland 变分原理于 I, 可以得到一列 $\{l_n\}\subset\Gamma$ 满足

$$c\leqslant I(l_n)<c+\frac{1}{n}, \tag{15.4}$$

$$I(l)>I(l_n)-\frac{1}{n}d(l,l_n),\quad l\neq l_n,\quad n=1,2,\cdots. \tag{15.5}$$

令

$$M(l)=\{t\in[0,1]\mid f\circ l(t)=I(l)\}.$$

显然 $M(l)$ 是非空紧集, 并且 $M\subset(0,1)$, 这是因为若 $t_0\in M(l)\cap\{0,1\}$, 则

$$f\circ l(t_0)=\max_{t\in[0,1]}f\circ l(t)\geqslant\inf_{\partial\Omega}f=\alpha,$$

而

$$f\circ l(t_0)\leqslant\max\{f(p_0),f(p_1)\}<\alpha,$$

这是一个矛盾.

记 $\Gamma_0=\{\psi\in C([0,1],X)\mid\psi(i)=\theta, i=0,1\}$, 它是 $C([0,1],X)$ 的一个线性闭子空间, 有模

$$\|\psi\|_{\Gamma_0}=\max_{t\in[0,1]}\|\psi(t)\|.$$

由 (15.4), $\forall h\in\Gamma_0$, $\|h\|_{\Gamma_0}=1$, $\forall\lambda_j\downarrow 0$, $\forall\xi_j\in M(l_n+\lambda_j h)$, 有

$$\lambda_j^{-1}\left[f\circ(l_n+\lambda_j h)(\xi_j)-f\circ l_n(\xi_j)\right]\geqslant-\frac{1}{n}.$$

由于 $\{\xi_j\}\subset[0,1]$, 故存在子列, 仍记作 ξ_j, 收敛于 η_n, 后者依赖于 l_n, λ_j 与 h, 取极限后有

$$df(l_n(\eta_n),h(\eta_n))\geqslant-\frac{1}{n}, \tag{15.6}$$

要想证明 $\exists\eta_n^*\in M(l_n)$ 使得

$$df(l_n(\eta_n^*),\varphi)\geqslant-\frac{1}{n},\quad\forall\varphi\in X,\ \|\varphi\|=1. \tag{15.7}$$

如其成立, 则令 $x_n = l_n(\eta_n^*)$, 即得

$$c \leqslant f(x_n) < c + \frac{1}{n},$$

$$\sup_{\|\varphi\|=1} |df(x_n, \varphi)| \leqslant \frac{1}{n}.$$

由 PS_c, $\{x_n\}$ 有收敛子列 $\{x_{n_j}\}, x_{n_j} \to x^*$, 即得 $f'(x^*) = 0$.

现在用反证法来证明 (15.7), 如不存在 η_n^* 使 (15.7) 成立, 则 $\forall \eta \in M(l_n)$, $\exists v_\eta \in X$, $\|v_\eta\| = 1$ 满足

$$df(l_n(\eta), v_\eta) < -\frac{1}{n},$$

便有 η 的一个邻域 $O_\eta \subset (0,1)$ 使得

$$df(l_n(\xi), v_\eta) < -\frac{1}{n}, \quad \forall \xi \in O_\eta,$$

注意到 $M(l_n)$ 是紧的, 它存在有限覆盖, 设 $M(l_n) \subset \bigcup_{i=1}^m O_{\eta_i}$, 对应着 $v\{v_{\eta_i}\}_1^m$, $\|v_{\eta_i}\| = 1$, 满足

$$df(l_n(\xi), v_{\eta_i}) < -\frac{1}{n}, \quad \forall \xi \in O_{\eta_i}, \ i = 1, \cdots, m, \tag{15.8}$$

构作 $\{O_{\eta_i}\}$ 的一个单位分解: $0 \leqslant \rho_i \leqslant 1, \operatorname{supp} \rho_i \subset O_{\eta_i}, 1 \leqslant i \leqslant m$, 满足

$$\sum_{i=1}^m \rho_i(\xi) = 1, \quad \forall \xi \in M(l_n),$$

令

$$v = v(\xi) = \sum_{i=1}^m \rho_i(\xi) v_{\eta_i}.$$

因为 $M(l_n) \subset (0,1)$, 所以 $v \in \Gamma_0$, $\|v\| \leqslant 1$. 事实上可以选取一个有限覆盖和一个 $\xi^* \in M(l_n)$ 使得只有一个 i_0 使 $\xi^* \in O_{\eta_i}$, 于是 $\|v\|_{\Gamma_0} = 1$, (15.8) 蕴含了

$$df(l_n(\xi), v(\xi)) < -\frac{1}{n}, \quad \forall \xi \in M(l_n),$$

它与 (15.6) 矛盾, 定理证毕. □

注 15.1 山路定理是由 A. Ambrosetti 和 P. H. Rabinowitz 在 1974 年以如上形式提出的. 它连同其各种推广和变形都已经被广泛地应用到各类变分问题中去了. 其原型则是 M. Morse 研究极小曲面多解时发现的墙定理 (Wall Theorem).

用 Ekeland 变分原理来证明山路定理是由史树中 (数学学报, 1, 1985, 348–355) 与 J.-P. Aubin 和 I. Ekeland (*Applied Nonlinear Analysis*, John Wiley and Sons, 1984) 分别独立提出来的. □

注 15.2 在以上定理中 Palais-Smale 条件起着关键的作用. 若其不成立, 则有下述 Brezis-Nirenberg 反例.

在 $\mathbb{R}^2$ 上考察函数

$$f(x,y) = x^2 + (1-x)^3 y^2,$$

令 $c = \inf_{x^2+y^2=1/4} f(x,y) > 0$, 它确有一个盆地: $\Omega = \{(x,y) \in \mathbb{R}^2 | \, f(x,y) \leqslant c\}$, 而且 $f(0,0) = 0, f(4,1) = -11$, 但直接验证 f 只有一个临界点 $(0,0)$. □

山路定理中的几何结构是更一般的环绕 (link) 结构的特殊情形.

定义 15.1 设 X 是一个 Banach 空间, $Q \subset X$ 是一个有边界 ∂Q 的紧流形, $S \subset X$ 是一个闭子集. 称 ∂Q 与 S 环绕, 如果

(1) $\partial Q \cap S = \varnothing$,

(2) 对任意连续的 $\varphi : Q \to X$ 满足: $\varphi|_{\partial Q} = id|_{\partial Q}$, 都有 $\varphi(Q) \cap S \neq \varnothing$.

环绕是两个集合在任意连续变化下的相交性质, 因此是一种拓扑性质.

例 15.1 在山路定理中, $Q = \{tx_0 + (1-t)x_1 \,|\, t \in [0,1]\}$, $S = \partial\Omega$, 而 $\partial Q = \{x_0, x_1\}$, 则 ∂Q 与 S 环绕. □

例 15.2 设 X 是一个 Banach 空间, X_1 是它的一个有限维线性子空间, X_2 是它的补空间, 即: $X = X_1 \oplus X_2$. 令

$$S = X_2, \quad Q = B_R \cap X_1,$$

其中 B_R 是 X 以 θ 为中心、$R > 0$ 为半径的闭球, 于是

$$\partial Q = \{x \in X_1 \,|\, \|x\| = R\}$$

(见图 15.2).

我们要证明 S 与 ∂Q 环绕. 正如前面所说的, 环绕是一个拓扑概念, 我们不得不使用拓扑工具. 一个比较简单的拓扑工具是 Brouwer 拓扑度.

在 $\mathbb{R}^n$ 上给定一个到自身的连续映射 f, 给定一个有界开集 $\Omega \subset \mathbb{R}^n$, 再给定一点 $p \notin f(\partial\Omega)$.

Brouwer 拓扑度 $\deg(f, \Omega, p)$ 是依赖于这三个变元 (f, Ω, p) 的取整数值的函数. 它具有如下重要性质.

1. (Kronecker 存在性) 若 $p \notin f(\partial\Omega)$, 且 $\deg(f, \Omega, p) \neq 0$, 则 $f^{-1}(p) \cap \Omega \neq \varnothing$.
2. (同伦不变性) 若 $F \in C([0,1] \times \bar{\Omega}, \mathbb{R}^n)$, $p \notin F([0,1] \times \partial\Omega)$, 则

$$\deg(F(t,\cdot), \Omega, p) = \text{const}.$$

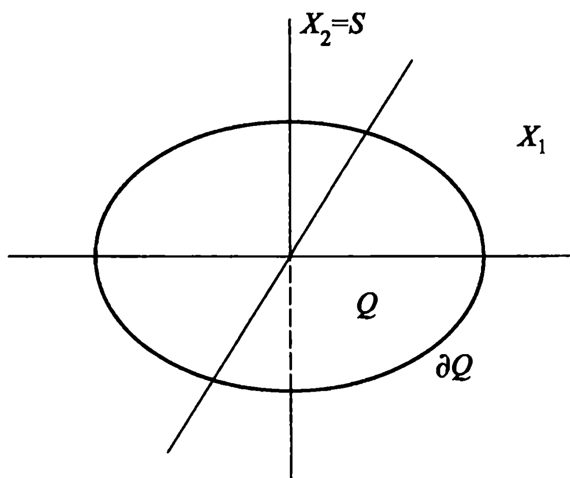


图 15.2 环绕之例 15.2

3. (区域可加性) 设 $\Omega_1, \Omega_2 \subset \Omega$ 都是有界开集, $\Omega_1 \cap \Omega_2 = \varnothing$, $p \notin f(\bar{\Omega} \backslash (\Omega_1 \cup \Omega_2))$, 则

$$\deg(f, \Omega, p) = \deg(f, \Omega_1, p) + \deg(f, \Omega_2, p).$$

4. (规范性)

$$\deg(f, \Omega, p) = \begin{cases} 1, & p \in \Omega, \\ 0, & p \notin \Omega. \end{cases}$$

5. (零点的代数和) 如果再设 $f \in C^1(\Omega, \mathbb{R}^n)$, 以及 $p \notin f(\partial\Omega)$ 是 f 的一个正则值, 即

$$\det(\partial_{x_i} f_j(x)) \neq 0, \quad \forall x \in f^{-1}(p),$$

则

$$\deg(f, \Omega, p) = \sum_{x_k \in f^{-1}(p) \cap \Omega} \operatorname{sgn} \det(\partial_{x_i} f_j(x_k)).$$ □

现在我们回到 ∂Q 与 S 环绕的证明. 显然有 $S \cap \partial Q = \varnothing$, 只要再证明 $\forall \varphi \in C(Q, X)$, 满足 $\varphi|_{\partial Q} = id|_{\partial Q}$ 都有

$$\varphi(Q) \cap S \neq \varnothing.$$

这等价于存在 $x_0 \in Q$ 使得

$$P \circ \varphi(x_0) = \theta,$$

其中 $P: X \to X_1$ 是一个投影算子. 为此我们定义 $F \in C([0,1] \times Q, \mathbb{R}^n)$,

$$F(t, x) = tP \circ \varphi(x) + (1-t)x.$$

因为

$$\theta \notin \partial Q = F(t, \partial Q), \quad \forall t \in [0,1],$$

由同伦不变性与规范性可得

$$\deg(P \circ \varphi, Q, \theta) = \deg(id, Q, \theta) = 1.$$

再利用 Kronecker 存在性就有 $x_0 \in Q$ 使得

$$P \circ \varphi(x_0) = \theta.$$

即得 S 与 ∂Q 环绕. □

例 15.3 设 X 是一个 Banach 空间, X_1 是它的一个有限维线性子空间, X_2 是它的补空间, 即 $X = X_1 \oplus X_2$. 又设 $e \in X_2, \|e\| = 1,\ R > \rho > 0$.

令

$$S = X_2 \cap \partial B_\rho(\theta),$$

$$Q = \{x_1 + te \,|\, (x_1, t) \in X_1 \times \mathbb{R}^1_+, \|x_1\|^2 + t^2 \leqslant R^2\},$$

从而

$$\partial Q = (B_R(\theta) \cap X_1) \cup (\partial B_R(\theta) \cap (X_1 \oplus \mathbb{R}^1 e))^+,$$

其中

$$(\partial B_R(\theta) \cap (X_1 \oplus \mathbb{R}^1 e))^+ = \{x_1 + te \,|\, (x_1, t) \in X_1 \times \mathbb{R}^1_+, \|x_1\|^2 + t^2 = R^2\}.$$

见图 15.3

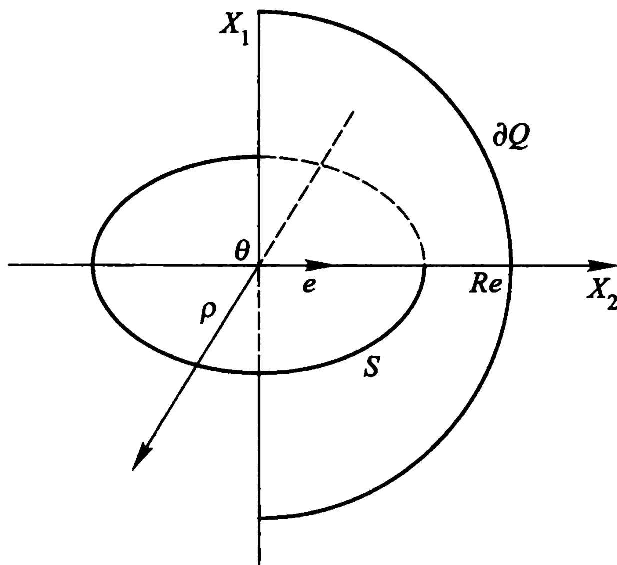


图 15.3 环绕之例 15.3

我们要证明 S 与 ∂Q 环绕. 显然还是有 $S \cap \partial Q = \varnothing$, 只要再证明对于满足 $\varphi|_{\partial Q} = id|_{\partial Q}$ 的任意 $\varphi \in C(Q, X)$ 都存在 $x_0 \in Q$, 使得

$$P \circ \varphi(x_0) = \theta,\ \|\varphi(x_0)\| = \rho,$$

其中 $P: X \to X_1$ 是投影算子. 这等价于证明

$$P \circ \varphi(x_1 + se) = \theta, \quad \|(I-P) \circ \varphi(x_1 + se)\| = \rho,$$

其中 $x_1 + se = x_0$.

为此作形变 $F \in C([0,1] \times Q, X_1 \times \mathbb{R}^1 e)$,

$$F(t, x_1+se) = [(1-t)x_1 + tP \circ \varphi(x_1+se)] + [(1-t)s + t\|(I-P)\circ\varphi(x_1+se)\| - \rho]e,$$

则

$$F(1, x_1 + se) = P \circ \varphi(x_1 + se) + [\|(I-P) \circ \varphi(x_1 + se)\| - \rho]e,$$

$$F(0, x_1 + se) = x_1 + (s - \rho)e.$$

并且当 $x_1 + se \in \partial Q$ 时,

$$F(t, x_1 + se) = x_1 + (s - \rho)e \neq \theta.$$

由同伦不变性得到

$$\deg(F(1,\cdot), Q, \theta) = \deg(F(0,\cdot), Q, \theta).$$

后者可以通过性质 5 (零点的代数和) 求出来:

$$F(0,\cdot)^{-1}(\theta, 0) = (\theta, \rho), \quad \det\left(\frac{\partial(x_1, s-\rho)}{\partial(x_1, s)}\right) = 1,$$

即得

$$\deg(F(1,\cdot), Q, \theta) = 1.$$

再由 Kronecker 存在性, 可知

$$F(1, x_1 + se) = (\theta, 0)$$

有解. 我们证明了 S 与 ∂Q 环绕. □

用与证明山路定理同样的步骤可以证明以下定理.

定理 15.2 设 X 是一个 Banach 空间, $Q \subset X$ 是一个有边界 ∂Q 的紧流形, ∂Q 与一个闭子集 $S \subset X$ 环绕, $f \in C^1(X, \mathbb{R}^1)$. 若存在 $\alpha < \beta$ 使得

$$\sup_{x \in \partial Q} f(x) \leqslant \alpha < \beta \leqslant \inf_{x \in S} f(x), \tag{15.9}$$

令

$$\Gamma = \{\varphi \in C(Q, X) \mid \varphi|_{\partial Q} = id|_{\partial Q}\}, \tag{15.10}$$

以及

$$c = \inf_{\varphi \in \Gamma} \max_{\xi \in Q} f(\varphi(\xi)). \tag{15.11}$$

如果 f 还满足 PS_c, 那么 $c\ (\geqslant \beta)$ 是 f 的一个临界值. □

§15.2 应用

山路定理和环绕定理在变分问题中有许多应用. 我们通过几个例子以窥见其一斑.

例 15.4 在实直线上给定一个以 $T > 0$ 为周期的连续函数 a. 定义位势函数

$$V(t, x) = -\frac{1}{2}|x|^2 + \frac{a(t)}{p+1}|x|^{p+1}. \tag{15.12}$$

假设 $p > 1,\ a(t) \geqslant \alpha > 0$.

求满足方程组

$$\ddot{x} + V_x(t, x) = 0 \tag{15.13}$$

的非平凡 T-周期解$x \in C^2([0, T], \mathbb{R}^N)$.

我们在空间 $H^1_{\text{per}}((0, T), \mathbb{R}^N) = \{x \in H^1((0, T), \mathbb{R}^N) \mid x(0) = x(T)\}$ 中定义泛函

$$I(x) = \int_0^T \left[\frac{1}{2}(|\dot{x}|^2 + |x|^2) - \frac{a(t)}{p+1}|x|^{p+1}\right] dt. \tag{15.14}$$

(15.13) 是它的 Euler-Lagrange 方程.

我们说明 $x = \theta$ 是 I 的一个局部极小点, 从而它是 (15.13) 的一个平凡解. 事实上,

$$I(\theta) = 0.$$

再由嵌入定理, $\|x\|_{p+1} \leqslant C\|x\|$, 其中 C 是一个常数, $\|x\| = \left(\int_0^T [|\dot{u}|^2 + |u|^2] dt\right)^{1/2}$. 因为 $\exists M > 0$ 使得 $|a(t)| \leqslant M, \forall t \in [0, T]$, 所以

$$(p+1)^{-1} \int_0^T a(t)|x(t)|^{p+1} dt \leqslant MC^{p+1}\|x\|^{p+1}.$$

取定 $\varepsilon > 0$ 足够小, 当 $x \in B_\varepsilon(\theta) \backslash \{\theta\}$ 时,

$$I(x) = \frac{1}{2}\|x\|^2 - (p+1)^{-1} \int_0^T a(t)|x(t)|^{p+1} dt \geqslant \frac{1}{4}\|x\|^2.$$

所以 $x=\theta$ 是一个局部极小点.

我们想用山路定理来得到一个非平凡的临界点. 为此取 "盆地" 外一低点: $e=\lambda\xi\sin\frac{2t\pi}{T}$, $\xi=(1,1,\cdots,1)$, 其中 $\lambda>0$ 充分大, 便有

$$I(e)=\lambda^2 n\left(\frac{2\pi^2}{T}+\frac{T}{2}\right)-\lambda^{p+1}n^{\frac{p+1}{2}}\int_0^T(p+1)^{-1}a(t)\left|\sin\frac{2t\pi}{T}\right|^{p+1}dt<0.$$

定义

$$\Gamma=\{i\in C([0,1],H_0^1((0,T),\mathbb{R}^N))\,|\,l(0)=\theta,l(1)=e\},$$

以及

$$c=\inf_{l\in\Gamma}\sup_{x\in l}I(x)\geqslant\frac{1}{4}\varepsilon^2.$$

只要能验证 PS_c, 就知道 c 一定是一个临界值.

设 $\{x_n\}$ 是一个 PS 序列,

$$\begin{cases}I(x_n)\to c,\\ \|I'(x_n)\|=\sup\limits_{\|\varphi\|=1}dI(x_n,\varphi)\to 0,\end{cases}$$

要证明 $\{x_n\}$ 含有收敛子列. 如今,

$$dI(x_n,\varphi)=\int_0^T[\dot{x_n}\dot{\varphi}+x_n\varphi-a(t)|x_n|^{p-1}x_n\varphi]\,dt\to 0,\ \forall\varphi\in H^1_{\mathrm{per}}((0,T),\mathbb{R}^N).\tag{15.15}$$

在 (15.15) 中取 $\varphi=x_n$, 就有

$$\int_0^T[|\dot{x}_n|^2+|x_n|^2-a(t)|x_n|^{p+1}]dt=o(\|x_n\|).\tag{15.16}$$

还有

$$\int_0^T\left[\frac{|\dot{x}_n|^2+|x_n|^2}{2}-a(t)\frac{|x_n|^{p+1}}{p+1}\right]dt\to c\quad(\neq 0).\tag{15.17}$$

比较 (15.16) 与 (15.17), 得到

$$\left(\frac{1}{2}-\frac{1}{p+1}\right)\int_0^T(|\dot{x}_n|^2+|x_n|^2)\,dt=C_1+o(\|x_n\|),$$

其中 C_1 是一个常数. 从而 $\|x_n\|$ 是有界的, 从而存在子列, 仍记作 x_n, 使得

$$x_n\rightharpoonup x_0,\ 于\ H^1_{\mathrm{per}}((0,T),\mathbb{R}^N).$$

最后要验证 $x_n \to x_0$, $H^1_{\text{per}}((0,T),\mathbb{R}^N)$.

事实上, 由嵌入定理,

$$x_n \to x_0 \quad (\text{于 } L^\infty([0,T],\mathbb{R}^N)).$$

再利用 PS 序列的假设

$$x_n - \left(-\frac{d^2}{dt^2}+1\right)^{-1}(a(t)|x_n|^{p-1}x_n) \to 0,\ H^1_{\text{per}}((0,T),\mathbb{R}^N).$$

可见 x_n 在 $H^1_{\text{per}}((0,T),\mathbb{R}^N)$ 是强收敛的. 即得 $x_n \to x_0$, 于 $H^1_{\text{per}}((0,T),\mathbb{R}^N)$.

我们验证了山路定理的一切条件, 所以这个问题有一个非平凡解 $x_0 \in H^1_{\text{per}}((0,T),\mathbb{R}^N)$. 再由正则性, $x_0 \in C^2_{\text{per}}((0,T),\mathbb{R}^N)$. □

注 15.3 同样的证明适用于

$$V(t,x) = \frac{-c}{2}|x|^2 + \frac{a(t)}{p+1}|x|^{p+1}$$

的情形, 其中 $c>0$. □

例 15.5 在例 15.4 中, 将位势函数改为

$$V(t,x) = \frac{c}{2}|x|^2 + \frac{a(t)}{p+1}|x|^{p+1}, \tag{15.12'}$$

其中 $c>0$, 而 $p>1$, $a(t) \geqslant \alpha > 0$. 求 (15.13) 的非平凡 T-周期解.

仍取空间 $H^1_{\text{per}}((0,T),\mathbb{R}^N)$, 相应的泛函改为

$$I(x) = \int_0^T \left[\frac{1}{2}(|\dot{x}|^2 - c|x|^2) - \frac{a(t)}{p+1}|x|^{p+1}\right] dt. \tag{15.14'}$$

容易发现 $x=\theta$ 仍是一个临界点, 但不再是极小点. 为了寻求非平凡的临界点, 我们来考察这个泛函水平集间的环绕集合. □

把方程 (15.13) 线性化, 对应的是线性方程

$$-\ddot{x}(t) = cx(t),\quad t\in[0,T],$$

连同周期边值条件.

我们先转向下列特征值问题:

$$-\ddot{x}(t) = \lambda x(t),\quad t\in[0,T], \tag{15.18}$$

其中 $x(0) = x(T)$, $\dot{x}(0) = \dot{x}(T)$.

按第 12 讲, 它有特征值

$$\lambda_k = \left(\frac{2k\pi}{T}\right)^2, \quad k = 0, 1, 2, \cdots$$

和特征函数

$$\cos\frac{2k\pi t}{T} \otimes e, \ \sin\frac{2k\pi t}{T} \otimes f,$$

其中 $e, f \in \mathbb{R}^N$, 当 $k \geqslant 1$; $e \in \mathbb{R}^N$, 当 $k = 0$. 对于每个 $k \geqslant 1$, 记

$$E_k = \operatorname{span}\left\{\cos\frac{2k\pi t}{T} \otimes e, \ \sin\frac{2k\pi t}{T} \otimes f \,\middle|\, e, f \in \mathbb{R}^N\right\}$$

为 $H^1_{\text{per}}((0,T), \mathbb{R}^N)$ 的 $2N$ 维线性子空间, 以及 $E_0 = \mathbb{R}^N$ 为其 N 维线性子空间.

我们可以把空间 $X = H^1_{\text{per}}((0,T), \mathbb{R}^N)$ 作直和分解: $X = X_1 \oplus X_2$,

$$X_1 = \bigoplus_{j=0}^{k} E_j, \tag{15.19}$$

其中 $k = \max\{j \in \mathbb{N} \mid \lambda_j \leqslant c\}$.

在子空间 X_2 上,

$$\int_0^T |\dot{x}(t)|^2 dt \geqslant \lambda_{k+1} \int_0^T |x(t)|^2 dt,$$

所以

$$I(x) \geqslant \frac{1}{2}\left(1 - \frac{c}{\lambda_{k+1}}\right) \int_0^T |\dot{x}(t)|^2 dt - \frac{1}{p+1} \int_0^T a(t)|x(t)|^{p+1} dt.$$

在例 15.4 中我们估计过

$$\int_0^T a(t)|x(t)|^{p+1} dt \leqslant MC^{p+1}\|x\|^{p+1} = o(\|x\|^2) \ \ (x \to \theta).$$

由于 $1 - \frac{c}{\lambda_{k+1}} > 0$, 所以存在 $\rho > 0$ 及 $\beta > 0$ 使得

$$I(x) \geqslant \beta \ \text{当} x \in \partial B_\rho(\theta) \cap X_2. \tag{15.20}$$

我们取 $S = \partial B_\rho(\theta) \cap X_2$.

在 X_1 上我们总有

$$I(x) \leqslant \int_0^T \left[\frac{1}{2}(|\dot{x}|^2 - c|x|^2)\right] dt \leqslant \int_0^T \left[\frac{1}{2}(\lambda_k - c)|x|^2\right] dt \leqslant 0.$$

现在我们取 $e = \cos\frac{2(k+1)\pi t}{T} \otimes e_1$, 其中 $e_1 = (1, 0, \cdots, 0) \in \mathbb{R}^N$. 注意到在任意有限维模空间上所有的模都是等价的, 在空间 $X_1 \oplus \mathbb{R}^1 e$ 上, 当 $\|x\| \to \infty$ 时, 一致地有

$$I(x) \leqslant (\lambda_{k+1} - c)\int_0^T |x|^2 dt - \frac{\alpha}{p+1}\int_0^T |x|^{p+1} dt \to -\infty.$$

现在我们取

$$Q = \{(x_1, t) \in X_1 \times \mathbb{R}^1_+ \mid \|x_1\|^2 + t^2 = R^2\}.$$

当 $R > \rho > 0$ 足够大时, 便有

$$I|_{\partial Q} \leqslant 0. \tag{15.21}$$

并且利用例 15.3, 还可见 S 与 ∂Q 环绕.

最后我们来验证 PS_c. 证明的步骤与例 15.5 是一样的, 只是在验证 PS 序列的 $H^1_{\mathrm{per}}((0,T), \mathbb{R}^N)$ 有界性时没有那么直接. 其证明如下. 设 $\{x_j\} \subset H^1_{\mathrm{per}}((0,T), \mathbb{R}^N)$ 满足 $I(x_j) \to c$, $I'(x_j) \to \theta$, 则有

$$\int_0^T [|\dot{x}_j(t)|^2 - c|x_j(t)|^2] dt = C_1 + o(\|x_j\|),$$

以及

$$\int_0^T a(t)|x_j(t)|^{p+1} dt = C_2 + o(\|x_j\|).$$

由 Hölder 不等式得

$$\int_0^T |x_j(t)|^2 dt \leqslant \left(\int_0^T a(t)|x_j(t)|^{p+1} dt\right)^{\frac{2}{p+1}} \left(\int_0^T a^{-\frac{2}{p-1}}(t) dt\right)^{\frac{p-1}{p+1}}$$

联合这三个不等式就得到

$$\int_0^T |\dot{x}_j(t)|^2 dt = C_3 + o(\|x_j\|),$$

从而 $\|x_j\| \leqslant C_4$.

现在根据 (15.20)、(15.21) 和定理 15.2, 我们证明了 I 有临界值 $c \geqslant \beta > 0$, 它对应着 (15.13) 的一个非平凡解. □

类似的方法可以应用于偏微分方程.

例 15.6 设 $\Omega \subset \mathbb{R}^n$ 是一个有界区域. 设 $1 < p < 2^* - 1, a \in C(\bar{\Omega}), a(x) \geqslant \alpha > 0$, $\forall x \in \Omega$. 求下列方程的弱解:

$$-\Delta u(x) = a(x)|u(x)|^{p-1}u(x), \tag{15.22}$$

其中 $u \in H_0^1(\Omega)$. 在空间 $H_0^1(\Omega)$ 中定义泛函

$$I(u) = \int_\Omega \left[\frac{1}{2}|\nabla u(x)|^2 - \frac{a(x)}{p+1}|u(x)|^{p+1}\right] dx, \tag{15.23}$$

(15.22) 是它的 Euler-Lagrange 方程.

显然 $u = \theta$ 是它的一个平凡解, 我们来寻求 I 的非平凡临界点. 利用 Poincaré 不等式,

$$I(u) \geqslant \int_\Omega \frac{1}{2}|\nabla u(x)|^2 dx - o(\|u\|^2) \geqslant \frac{1}{4}\int_\Omega |\nabla u(x)|^2 dx \quad (\|u\| \to 0),$$

即当 $r > 0$ 充分小时,

$$I|_{\partial B_r(\theta)} \geqslant \frac{r^2}{4}.$$

现在任意选取一个非零函数 $\varphi \in H_0^1(\Omega)$,

$$I(t\varphi) = t^2 \int_\Omega \frac{1}{2}|\nabla\varphi(x)|^2 dx - t^{p+1}\int_\Omega \frac{a(x)}{p+1}|\varphi(x)|^{p+1} dx \to -\infty.$$

我们取 $p_0 = \theta, p_1 = t\varphi$, t 足够大时, 就得到山路定理的几何结构.

剩下来就只要验证 PS_c 了. 设序列 u_j 满足: $I(u_j) \to c$, $I'(u_j) \to \theta$ $(H_0^1(\Omega))$, 要证明它有收敛子序列. 事实上, 我们有

$$\int_\Omega \left[\frac{1}{2}|\nabla u_j(x)|^2 - \frac{a(x)}{p+1}|u_j(x)|^{p+1}\right] dx \to c, \tag{15.24}$$

以及

$$\int_\Omega [\nabla u_j(x)\varphi(x) - a(x)|u_j(x)|^{p-1}u_j(x)\varphi(x)]dx = o(\|\varphi\|), \ \forall\ \varphi \in H_0^1(\Omega). \tag{15.25}$$

将 $\varphi = u_j$ 代入得到

$$\int_\Omega [|\nabla u_j(x)|^2 - a(x)|u_j(x)|^{p+1}]dx = o(\|u_j\|). \tag{15.26}$$

联合 (15.24) 与 (15.26) 得到

$$\left(\frac{1}{2} - \frac{1}{p+1}\right)\int_\Omega |\nabla u_j(x)|^2 dx = C + o(\|u_i\|).$$

所以 u_j 是 $H_0^1(\Omega)$ 中的有界列, 于是有弱收敛子列 $u_{j'} \rightharpoonup u_0$.

我们来证明它们是强收敛：$u_{j'} \to u_0$, 于 $H_0^1(\Omega)$. 由 (15.25) 可见

$$u_{j'} = (-\Delta)^{-1}(a|u_{j'}|^{p-1}u_{j'} + o(\|u_{j'}\|)).$$

我们注意到

(1) $H_0^1(\Omega) \hookrightarrow L^{p+1}(\Omega)$ 是紧的,

(2) $u \mapsto a|u|^{p-1}u$ 是 $L^{p+1}(\Omega) \to L^{\frac{p+1}{p}}(\Omega) \hookrightarrow (H_0^1)^*$ 有界连续的,

(3) $(-\Delta)^{-1} \in L((H_0^1)^*, H_0^1)$ 是连续的,

所以有强收敛 $u_{j'} \to u_0$.

我们完成了 PS_c 的验证. 应用山路定理, 得到临界值 $c \geqslant r^2/4 > 0$, 所以对应的是一个非平凡临界点.

当 $c < 0$ 时, 方程

$$-\Delta u(x) = cu(x) + a(x)|u(x)|^{p-1}u(x) \tag{15.22$'$}$$

与方程 (15.22) 有同样的结论, 证明是一样的.

当 $c > 0$ 时用环绕定理也可以证明方程 (15.22′) 有非平凡解. □

注 15.3 山路定理和它的推广 ——环绕定理都是更一般的极小极大原理的特殊情形. 极小极大原理最早是由 G.D. Birkhoff 在研究闭测地线时提出来的. 经 L. Liusternik, L. Schnirelmann 以及 M. Krasnoselski 等人的发展, 成为临界点理论的重要组成部分. 到了 20 世纪 60 年代, R.S. Palais 将 Liusternik-Schnirelmann 理论推广到了无穷维流形. 自从 A. Ambrosetti 和 P.H. Rabinowitz 提出山路定理以来, 极小极大原理得到了更多的发展与应用. 与此平行发展的临界点理论的另一个分支是 Morse 理论, 它们共同组成大范围变分学或拓扑变分方法. 因为涉及的内容很多, 已远远超出本课程的范围, 本书不再介绍. 有兴趣的读者可以阅读有关参考书籍, 如 [Ch], [MW], [St], [Ra] 等. □

第十六讲　周期解、异宿轨与同宿轨

§16.1　问题

我们从单摆的自由运动出发 (见图 16.1), 介绍周期解、第二周期解与异宿轨的概念. 质量为 m, 摆长为 l 的小球受重力来回摆动. 记摆的张角为 φ, 那么动能是 $\frac{1}{2}mgl^2$, 位能是 $mgl(1-\cos\varphi)$. Lagrange 函数是 $\frac{1}{2}mgl^2 - mgl(1-\cos\varphi)$. 动力学方程是

$$\ddot{\varphi} + \alpha \sin\varphi = 0,$$

其中 $\alpha = \omega_0^2 = g/l$.

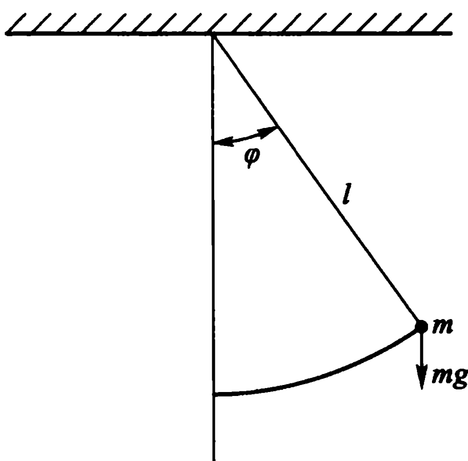


图 16.1　单摆的运动

在相平面 $(x,y)=(\varphi,\dot{\varphi})$ 上来讨论 (见图 16.2),

$$\begin{cases} \dot{x}=y, \\ \dot{y}=-\alpha\sin x, \end{cases}$$

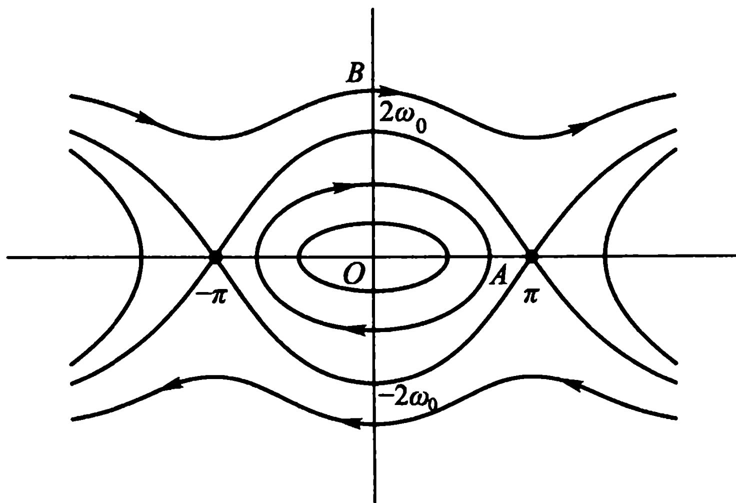


图 16.2 相图

能量是

$$E=\frac{m}{2}l^2y^2+mgl(1-\cos x).$$

图上画了能量的等高线和平衡点: $(0,0)$, $(\pm\pi,0)$.

当 $A\in(0,\pi)$ 时, 过初始点 $(x,y)=(A,0)$ 的能量是 $E=mgl(1-\cos A)$. 运动方程是

$$y=\pm\omega_0\sqrt{2(\cos x-\cos A)}.$$

摆以 $T(A)=\frac{2\sqrt{2}}{\omega_0}\int_0^A\frac{dx}{\sqrt{\cos x-\cos A}}$ 为周期作周期运动.

$$\lim_{A\to 0}T(A)=\frac{2\pi}{\omega_0},\quad \lim_{A\to\pi-0}T(A)=\infty.$$

在上、下半平面过 $(0,\pm 2\omega_0)$ 连接 $(-\pi,0)$ 与 $(\pi,0)$ 的轨道: $(x(t),\dot{x}(t))\to(\pm\pi,0)(t\to\pm\infty)$ 称为异宿轨.

还有, 设 $|B|>2\omega_0$, 过初始点 $(x,y)=(0,\pm B)$ 的能量是 $E=ml^2B^2/2$ 的解,

$$y=\pm\sqrt{B^2-2\omega_0^2(1-\cos x)}$$

分别为上、下半平面的周期曲线, 令

$$t(x)=\int_0^x\frac{ds}{\sqrt{B^2-2\omega_0^2(1-\cos s)}},$$

以及

$$\tau(B) = t(2\pi),$$

则

$$p(x) = t(x) - \frac{\tau(B)}{2\pi}$$

是一个 2π 周期的函数. 我们有

$$\begin{cases} x(t+\tau(B)) = x(t) + 2\pi, \\ y(t+\tau(B)) = y(t). \end{cases}$$

在这个意义上, 我们称之为第二类周期解.

在动力系统中连接鞍点平衡点 p 到自身的轨道: $x(t) \to p$ $(t \to \pm\infty)$ 称为同宿轨 (见图 16.3).

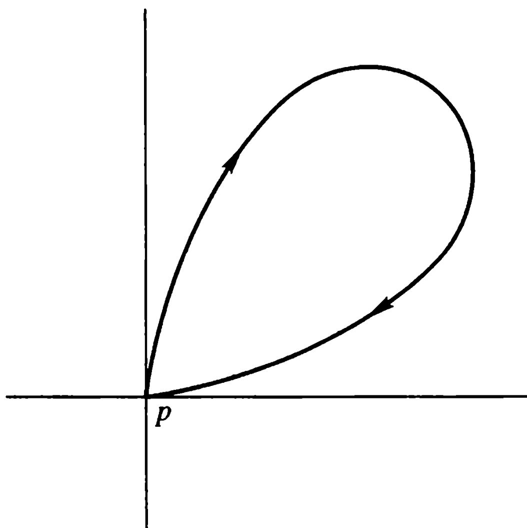


图 16.3

§16.2 周期解

在非线性振动中, 人们关心下列方程的周期解:

$$\ddot{u} + \nabla_u V(t,u) = 0, \ \ u(0) = u(T), \ \dot{u}(0) = \dot{u}(T). \tag{16.1}$$

例如, 月亮的引力 (潮汐力) 产生的重力加速度 g 是一个 T-周期函数, 位能

$$V(t,u) = \frac{g(t)}{2\pi} \cos 2\pi u.$$

一般来说, 设 $V \in C^1([0,T] \times \mathbb{R}^N, \mathbb{R}^1)$. 引入 Lagrange 函数

$$L(t,u,p) = \frac{|p|^2}{2} - V(t,u),$$

对应泛函是

$$I(u) = \int_0^T \left[\frac{|\dot{u}|^2}{2} - V(t,u)\right] \mathrm{d}t.$$

我们把 (16.1) 看成 I 的 E-L 方程, 并取空间

$$H^1_{\text{per}}([0,T], \mathbb{R}^N) = \{u \in H^1([0,T], \mathbb{R}^N) \,|\, u(0) = u(T)\}.$$

有如下简单的情形.

I. 假设 $u \mapsto -V(t,u)$ 是连续的、凸的, 并且

$$F(u) = -\int_0^T V(t,u)\,dt \to +\infty \quad (\|u\|_{\mathbb{R}^N} \to \infty). \tag{16.2}$$

例如, $N = 1, V(t,u) = -|u|^p(1 + \varepsilon \sin t)$, 其中 $p > 1$. 因为

$$-V_{uu} = -p(p-1)|u|^{p-2}(1 + \varepsilon \sin t) > 0.$$

这时 I 是凸的. I 是弱下半连续的, 从而是序列弱下半连续的. 接下来验证: I 是下方有界并且是强制的.

由于 $u \to -V(t,u)$ 是连续凸的, F 是 $\mathbb{R}^N$ 上的连续凸函数. 于是 $\exists\, x_0 \in \mathbb{R}^N$, 使其达到极小值. 我们有

$$0 = F'(x_0) = \int_0^T V_u(t,x_0)\,dt. \tag{16.3}$$

利用 $-V$ 的凸性, 有不等式

$$-V(t,u) \geqslant -(V(t,x_0) + V_u(t,x_0)(u - x_0)). \tag{16.4}$$

现在 $\forall\, u \in H^1_T(0,T)$, 作分解 $u = \tilde{u} + \bar{u}$, 其中 $\bar{u} = \frac{1}{T}\int_0^T u(t)dt$. 联合 (16.3) 和 (16.4) 得到

$$\begin{aligned} -\int_0^T V(t,u(t))dt &\geqslant -\int_0^T (V(t,x_0) + V_u(t,x_0)(u(t) - x_0))dt \\ &= -\int_0^T (V(t,x_0) + V_u(t,x_0)(u(t) - \bar{u}))dt. \end{aligned} \tag{16.5}$$

令

$$c_1 = -\int_0^T V(t, x_0)dt, \quad c_2 = \left(\int_0^T |V_u(t, x_0)|^2 \, dt\right)^{1/2},$$

则

$$\begin{aligned} I(u) &\geqslant \frac{1}{2}\int_0^T |\dot{\tilde{u}}|^2 dt - \int_0^T [V(t, x_0) + V_u(t, x_0)\tilde{u}(t)]dt \\ &\geqslant \frac{1}{2}\int_0^T |\dot{\tilde{u}}|^2 dt - c_1 - c_2 \left(\int_0^T |\tilde{u}|^2 dt\right)^{1/2} \end{aligned}$$

利用 Wirtinger 不等式, $\int_0^T |\tilde{u}|^2 dt \leqslant c_3^2 \int_0^T |\dot{\tilde{u}}|^2 dt$, 即得

$$\begin{aligned} I(u) &\geqslant \frac{1}{2}\int_0^T |\dot{\tilde{u}}|^2 dt - c_1 - c_2 c_3 \left(\int |\dot{\tilde{u}}|^2 dt\right)^{1/2} \\ &\geqslant \frac{1}{4}\int_0^T |\dot{\tilde{u}}|^2 dt - c_4. \end{aligned}$$

由此可见, I 是否为强制的, 就看 $\bar{u}$ 能否被 $I(u)$ 从上方控制. 注意到

$$\begin{aligned} -V\left(t, \frac{\bar{u}}{2}\right) &= -V\left(t, \frac{u(t) - \tilde{u}(t)}{2}\right) \\ &\leqslant -\frac{1}{2}(V(t, u(t)) + V(t, -\tilde{u}(t))). \end{aligned}$$

推得

$$I(u) \geqslant \frac{1}{2}\int_0^T |\dot{\tilde{u}}|^2 dt - 2\int_0^T V\left(t, \frac{\bar{u}}{2}\right) dt + \int_0^T V(t, -\tilde{u}(t))dt.$$

由于 $\|\tilde{u}\|_\infty \leqslant C\|\tilde{u}\|_{H^1}$ 能被 $I(u)$ 所控制, 所以 $-\int_0^T V(t, -\tilde{u}(t))dt$ 能被 $I(u)$ 所控制, 推得

$$-\int_0^T V\left(t, \frac{\bar{u}}{2}\right) dt \leqslant c_5 I(u) + c_6.$$

利用 (16.2), 可见 $\bar{u}$ 必被 $I(u)$ 所控制. 这就证明了 I 是强制的.

由此可见方程 (16.1) 有解 $u \in H^1_{\text{per}}([0, T])$. 再由正则性 $u \in C^2$.

最后, 我们来看周期条件. 从嵌入定理以及 $u \in H^1_{\text{per}}([0, T])$ 可见 $u(0) = u(T)$. 剩下还要证明 $\dot{u}(0) = \dot{u}(T)$. 在积分形式的 E-L 方程中取以 T 为周期的

$\varphi \in C^\infty([0,T])$, 得到

$$0 = \int_0^T [\dot{u}\dot{\varphi} - V_u(t,u)\varphi]dt$$
$$= \int_0^T [-\ddot{u} - V_u(t,u)]\varphi dt + \dot{u}\varphi|_0^T$$
$$= \dot{u}(T)\varphi(T) - \dot{u}(0)\varphi(0).$$

由于 $\varphi(0) = \varphi(T)$ 是任意的, 我们就证明了

$$\dot{u}(T) = \dot{u}(0).$$

定理 16.1 在 (16.2) 的假设下, 方程 (16.1) 存在一个 C^2 的周期解. □

II. V 是连续的、周期的. 假设存在线性无关的 $e_1, \cdots, e_N \in \mathbb{R}^N$ 使得

$$V(t, u+e_i) = V(t,u), \quad \forall (t,u) \in [0,T] \times \mathbb{R}^N. \tag{16.6}$$

我们还是要验证

$$I(u) = \int_0^T \left(\frac{1}{2}|\dot{u}(t)|^2 + V(t,u(t))\right) dt$$

的序列弱下半连续性和强制性.

按第十一讲的 Morrey 定理, I 当然是序列弱下半连续的.

由于 V 是连续的又满足 (16.6), 所以存在常数 C 使得

$$|V(t,u)| \leqslant C.$$

从

$$I(u) \geqslant \frac{1}{2}\int_0^T |\dot{u}|^2 dt - CT$$

还不能直接推出 I 的强制性. 为此作分解

$$u = \tilde{u} + \bar{u},$$

其中 $\bar{u} = \frac{1}{T}\int_0^T u(t)dt$. 记

$$X = \{u \in H^1_{\text{per}}([0,T]) \,|\, \bar{u} = 0\}.$$

X 是 $H^1_{\text{per}}([0,T])$ 上的闭线性子空间.

再由 Wirtinger 不等式,

$$\int_0^T |\dot{u}|^2 dt = \int_0^T |\dot{\tilde{u}}|^2 dt \geqslant \alpha \int_0^T |\tilde{u}|^2 dt,$$

即得 I 在 X 上是强制的.

设 $\{u_j\}$ 是 I 的一串极小化序列, 作分解 $u_j = \tilde{u}_j + \bar{u}_j$, 则 $\{\tilde{u}_j\}$ 有子列弱收敛. 注意到 V 的周期性 $V\left(u + \sum_{i=1}^N \lambda_i e_i\right) = V(u)$, 有

$$I\left(u + \sum_{i=1}^N \lambda_i e_i\right) = I(u), \quad \forall (\lambda_1, \cdots, \lambda_n) \in \mathbb{Z}^n.$$

虽然 $\{\bar{u}_j\}$ 可能无界, 但去掉整数部分以后就是有界的了, 即 $\exists (\lambda_1^{(j)}, \cdots, \lambda_n^{(j)}) \in \mathbb{Z}^n$, 使得

$$\|\bar{u}_j + \sum_{i=1}^N \lambda_i^{(j)} e_i\|_{\mathbb{R}^n} \leqslant \sum_{i=1} \|e_i\| \triangleq A.$$

定义 $v_j = \tilde{u}_j + \left(\bar{u}_j + \sum_{i=1}^N \lambda_i^{(j)} e_i\right)$, 则 $I(v_j) = I(u_j)$, 如今 v_j 是一个有界的极小化列, 有弱收敛的子列

$$v_j \rightharpoonup u^*.$$

于是按同样方法可证明 u^* 是极小点. 再用上面同样步骤验证 u^* 是周期函数. 由正则性, $u \in C^2(\mathbb{R}^1)$.

定理 16.2 在 (16.6) 的假设下, 方程 (16.1) 存在 C^2 的周期解. □

III. 环面上的周期解.

用与上面完全相同的方法, 我们可以研究在环面 $T^N = \mathbb{R}^N/\mathbb{Z}^N$ 上泛函 E-L 方程的周期解. 设在环面上给定 Lagrange 函数

$$L(t, u, p) : \mathbb{R}^1 \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}^1,$$

满足

$$L(t + \mathbb{Z}, u + \mathbb{Z}^n, p) = L(t, u, p),$$

以及下列条件:

$$\begin{cases} c^{-1} \leqslant L_{pp} \leqslant c, \\ |L_{pt}| + |L_{pu}| \leqslant c(1 + |p|), \\ |L_u| \leqslant c(1 + |p|^2), \end{cases}$$

则泛函

$$I(u)=\int_0^1 L(t,u(t),\dot{u}(t))dt$$

的 E-L 方程

$$\int_0^1 \{L_p(t,u(t),\dot{u}(t))\dot{\phi}(t)+L_u(t,u(t),\dot{u}(t))\phi(t)\}dt=0,\quad \forall\,\phi(t)\in H^1_{\text{per}}([0,1],\mathbb{R}^N)$$

存在周期解 $u\in H^1_{\text{per}}([0,1],\mathbb{R}^N)$. 再由正则性, $u\in C^2([0,1],\mathbb{R}^N)$. □

IV. 环面上的第二类周期解 $M_{q,p}$.

在与前面相同的条件下, 我们来考察第二类周期解.

$\forall\,(q,p)\in\mathbb{Z}^2, q\neq 0$, u 称为 (q,p) 型周期解, 如果

$$u(t+q)=u(t)+p.$$

我们在摆的自由振动中所说的周期解就是 $(1,0)$ 型周期解, 所说的第二类周期解就是 $(1,1)$ 型周期解.

在 $M_{q,p}=\left\{\left\{\frac{p}{q}t\right\}\,\middle|\,t\in\mathbb{R}^1\right\}+W^{1,2}_{\text{per}}([0,q])$ 中, 求得

$$I(u)=\int_0^q L(t,u(t),\dot{u}(t))\,\mathrm{d}t$$

的极小点就是所要的 (q,p) 型周期解.

采用同样的极小化方法, 以及同样的论证就能验证 I 的极小点是存在的. □

§16.3 异宿轨

由定义, 异宿轨是连接向量场（非退化）零点间的轨道. 例如给定 Lagrange 函数

$$L(t,u,p)=\frac{1}{2}|p|^2-V(t,u),$$

假设

(1) $V\in C^1(\mathbb{R}^N,\mathbb{R}^1)$,

(2) V 对 t 是 T 周期的,

(3) $V(t,u)\leqslant 0$, 只有两个非退化的极大点 θ 与 ξ.

$$V(t,\theta)=V(t,\xi)=0,\quad V_u(t,\theta)=V_u(t,\xi)=0,$$

其中 $V_{uu}(t,\theta)$ 与 $V_{uu}(t,\xi)$ 都是负定的,

(4) $\exists V_0 < 0$, 使得

$$\varlimsup_{|u|\to\infty} V(t,u) \leqslant V_0.$$

求连接 θ 与 ξ 下列方程组的解:

$$\begin{cases} \ddot{u} + V_u(t,u) = 0, t \in \mathbb{R}^1, \\ \dot{u}(-\infty) = \dot{u}(+\infty) = 0, u(-\infty) = \theta,\ u(+\infty) = \xi. \end{cases} \tag{16.7}$$

我们采用变分方法求解. 取空间

$$\hat{E} = \left\{ u \in H^1_{\mathrm{loc}}(\mathbb{R}^1, \mathbb{R}^N) \,\middle|\, \int_{\mathbb{R}^1} |\dot{u}|^2 < \infty \right\},$$

有模

$$\|u\|^2_{\hat{E}} = \int_{\mathbb{R}^1} |\dot{u}|^2 dt + |u(0)|^2.$$

定义泛函

$$I(u) = \int_{-\infty}^{\infty} \left(\frac{1}{2}|\dot{u}|^2 - V(t,u) \right) dt.$$

I 在 $\hat{E}$ 上可以不取有限值, 但因我们只关心极小化序列 $\{u_j\}$, 所以只涉及 I 取有限值的集合. 现在我们定义

$$\Gamma(\theta,\xi) = \{ u \in \hat{E} \,|\, u(-\infty) = \theta, u(+\infty) = \xi \},$$

其中 $u(\pm\infty) = \lim_{t\to\pm\infty} u(t)$. 求

$$\min_{u \in \Gamma(\theta,\xi)} I(u),$$

I 是序列弱下半连续的. 这可以从在任意有限区间上应用 Morrey 定理, 再取极限得到.

因为 $I \geqslant 0$, 所以它是下方有界的.

1° 我们要验证 I 的强制性. 由

$$\frac{1}{2}\int_{\mathbb{R}^1} |\dot{u}|^2 dt \leqslant I(u) \leqslant C$$

可知我们只要能控制 $|u(0)|$ 就够了.

当 $u \in \Gamma(\theta,\xi)$ 时, 我们有估计

$$|u(b) - u(a)| \leqslant \int_a^b |\dot{u}(t)|\, dt \leqslant (b-a)^{1/2} \left(\int_a^b |\dot{u}(t)|^2\, dt \right)^{1/2}$$

只要

$$-V(t,u(t)) \geqslant M_1, \tag{16.8}$$

就有

$$\begin{aligned} I(u) &\geqslant \int_a^b \left(\frac{1}{2}|\dot{u}|^2 - V(t,u)\right) \mathrm{d}t \\ &\geqslant \frac{1}{2}\frac{|u(b)-u(a)|^2}{|b-a|} + M_1|b-a| \\ &\geqslant \sqrt{2M_1}|u(b)-u(a)| \end{aligned} \tag{16.9}$$

然而由假设 (3) 与 (4), 取定 $\varepsilon > 0$, 就存在 $M_1 > 0$, 当

$$u(t) \notin B_\varepsilon(\theta) \cup B_\varepsilon(\xi)$$

时, (16.8) 成立.

注意到 V 对 t 是 T-周期的, 令

$$(\tau_i u)(t) = u(t - iT), \quad \forall\, i \in \mathbb{Z},$$

则 $I(\tau_i u) = I(u)$.

为了得到 $|u(0)|$ 的控制, 我们利用 u 的 τ_i 不变性. $\forall\, u \in \Gamma(\theta,\xi), \forall\, \varepsilon > 0, \exists\, i_0$ 使得 $\tau = \tau_{i_0}$ 满足

$$\tau u(t) \in B_\varepsilon(\theta), \quad \forall t < 0, \quad \tau u(0) \in \partial B_\varepsilon(\theta),$$

用 $\tilde{u} = \tau u$ 替换 u, 就有 $\|\tilde{u}\| = \varepsilon$. 我们证明了 I 是强制的. 以后, 对于 $\Gamma(\theta,\xi)$ 中的任意 u 都用 $\widetilde{u}$ 替换, 有 $I(u) = I(\widetilde{u})$.

2° 再验证 $\Gamma(\theta,\xi)$ 对极小化序列是弱闭的. 即要证明: 若 $\{u_j\} \subset \Gamma(\theta,\xi)$, $u_j \rightharpoonup u$, 并且 $I(u_j) \to \inf_{\Gamma(\theta,\xi)} I$, 则 $u \in \Gamma(\theta,\xi)$.

从 u 是极小化序列的弱极限以及 I 的弱下半连续性立即可得

$$I(u) \leqslant \inf_{\Gamma(\theta,\xi)} I.$$

利用 (16.9), $u \in L^\infty(\mathbb{R}^1,\mathbb{R}^N)$, 所以有 ω- 极限点, 即 $\exists\, t_i \to +\infty, \exists\, \alpha \in \mathbb{R}^N$ 使得 $u(t_i) \to \alpha$.

1) 我们说明极限点只有一个. 用反证法. 倘若有 $t_i' \to +\infty$, 使得 $u(t_i') \to \beta$, 那么因为 $I(u) < \infty$, 所以从 (16.9) 得到

$$|u(t_i) - u(t_i')| \leqslant \frac{1}{\sqrt{2M_1}}\left|\int_{t_i}^{t_i'} \left[\frac{1}{2}|\dot{u}|^2 - V(t,u)\right] dt\right| \to 0.$$

即得 $\alpha=\beta$.

2) 证明 $\alpha=\theta$ 或 ξ. 用反证法. $\exists t_1>0$, 使得当 $t>t_1$ 时,

$$u(t)\notin B_\varepsilon(\theta)\cup B_\varepsilon(\xi).$$

于是

$$I(u)\geqslant-\int_{t_1}^{\infty}V(t,u(t))\mathrm{d}t\geqslant\int_{t_1}^{\infty}M_1dt=\infty.$$

矛盾.

3) 证明 $\alpha=\xi$. 如其不然, $\alpha=\theta$. 那么 $\forall\delta>0,\exists t_\delta>0$, 当 $t\geqslant t_\delta$ 时, $u(t)\in B_\delta(\theta)$.

因为 $u_j\rightharpoonup u\in\hat{E}$, 而且当 $t\leqslant 0$ 时, $u_j(t)\in B_\varepsilon(\theta)$, $u_j(0)\in\partial B_\varepsilon(\theta)$.

取 $\delta<\varepsilon/4$ 以及 $t_1>t_\delta+1$, 使得 $u(t_1)\in B_\delta(\theta)$.

$\exists j_0$, 当 $j\geqslant j_0$ 时, $\|u_j-u\|_{L^\infty([0,t_1])}<\delta$, 从而 $u_j(t_1)\in B_{2\delta}(\theta)$. 由 (16.9),

$$I(u_j)\geqslant\int_{t_1}^{\infty}\left[\frac{1}{2}|\dot{u}_j|^2-V(t,u_j)\right]dt+\frac{\varepsilon}{2}\sqrt{2M_1}.$$

构造序列

$$v_j(t)=\begin{cases}\theta, & t<t_1-1,\\ (t-t_1+1)u_j(t_1), & t\in[t_1-1,t_1],\\ u_j(t), & t>t_1,\end{cases}$$

则 $v_j\in\Gamma(\theta,\xi)$, 并且

$$I(v_j)=\int_{t_1-1}^{t_1}\left[\frac{1}{2}|u_j(t_1)|^2-V(t,v_j(t))\right]dt+\int_{t_1}^{\infty}\left[\frac{1}{2}|\dot{u}_j(t)|^2-V(t,u_j(t))\right]dt,$$

而

$$\int_{t_1}^{\infty}\left[\frac{1}{2}|\dot{u}_j(t)|^2-V(t,u_j(t))\right]dt\leqslant I(u_j)-\frac{\varepsilon}{2}\sqrt{2M_1},$$

$$\int_{t_1-1}^{t_1}\left[\frac{1}{2}|u_j(t_1)|^2-V(t,v_j(t))\right]dt\leqslant 2\delta^2+\max_{|u|\leqslant 2\delta}(-V(t,u)).$$

再取 $\delta>0$, 使得

$$2\delta^2+\max_{|u|\leqslant 2\delta}(-V(t,u))<\frac{\varepsilon}{4}\sqrt{2M_1}.$$

如此得到

$$I(v_j)\leqslant I(u_j)-\frac{\varepsilon}{4}\sqrt{2M_1},$$

但 $I(u_j)\to\inf_{\Gamma(\theta,\xi)}I$, 这与 $\{v_j\}\subset\Gamma(\theta,\xi)$ 矛盾!

现在我们证明了 $\Gamma(\theta,\xi)$ 对极小化序列是弱闭的, 从而泛函 I 在其上存在极小点, 记作 u^*. 再由正则性, $u^*\in C^2(\mathbb{R}^N)$.

3° 最后验证 $\dot{u}^*(\pm\infty)=0$.

已知当 $t>t_1$ 时, $u^*(t)\in B_\varepsilon(\xi)$, 从而存在 $\beta_1>0,\beta_2>0$ 使得

$$-V(t,u^*(t))\geqslant \beta_1|u^*(t)-\xi|^2,$$
$$|V_u(t,u^*(t))|\leqslant \beta_2|u^*(t)-\xi|.$$

由假设 (3),

$$\beta_1\int_{t_1}^{\infty}|u^*(t)-\xi|^2dt\leqslant -\int_{t_1}^{\infty}V(t,u^*(t))dt\leqslant I(u^*),$$

以及

$$\int_{t_1}^{\infty}|\ddot{u}^*|^2dt=\int_{t_1}^{\infty}|V_u(t,u^*(t))|^2dt\leqslant \beta_2^2\int_{t_1}^{\infty}|u^*(t)-\xi|^2dt,$$

推出 $\int_{t_1}^{\infty}|\ddot{u}^*|^2dt<\infty$. 连同 $\dot{u}^*\in L^2(\mathbb{R}^1)$, 就有 $\dot{u}^*(+\infty)=0$. 同理证明 $\dot{u}^*(-\infty)=0$.

定理 16.3 *在假设 (1) − (4) 之下, 方程 (16.7) 存在异宿轨.*

§16.4　同宿轨

在直线上给定一个以 $2T>0$ 为周期的连续函数 $a\in C(\mathbb{R}^1)$. 假设 $\mu>2$, $\exists\alpha>0$ 使得 $a(t)\geqslant\alpha$. 定义位势函数

$$V(t,x)=-\frac{1}{2}|x|^2+a(t)|x|^\mu. \tag{16.10}$$

我们要寻求一条从 $x=0$ 出发的满足方程

$$\ddot{x}+V_x(t,x)=0 \tag{16.11}$$

的同宿轨: $x\in H^1(\mathbb{R}^1,\mathbb{R}^1)$.

现在采用的办法是 $\forall k\in\mathbb{N}$, 求方程 (16.11) 的一个 $2kT$-周期解 x_k. 再让 $k\to\infty$, 看这串解 x_k 的极限是否存在? 是否还满足方程 (16.11)? 是否同宿轨?

为此, $\forall k\in\mathbb{N}$, 引入空间 $X_k=H^1_{\text{per}}([-kT,kT])$, 有模

$$\|x\|_k^2=\int_{-kT}^{kT}\left[|\dot{x}(t)|^2+|x(t)|^2\right]dt.$$

定义泛函

$$I_k(x) = \frac{1}{2}\|x\|_k^2 - \int_{-kT}^{kT} a(t)|x(t)|^\mu dt.$$

由第十五讲例 15.5 可见, I_k 具有山路引理的几何结构:

$$I_k(0) = 0, \quad I_k(x) = \frac{1}{2}\|x\|_k^2 + o(\|x\|_k^2), \quad \exists\, \varphi \in X_k \backslash \{0\}, \text{使得 } I(t\varphi) \to -\infty.$$

I_k 还满足 Palais-Smale 条件, 从而它存在一个临界点 $x_k \in X_k$, 即

$$\int_{-kT}^{kT} [\dot{x}_k \dot{\varphi} + x_k \varphi - \mu a |x_k|^{\mu-2} x_k \varphi] dt = 0, \quad \forall\, \varphi \in X_k, \tag{16.11}$$

以及

$$c_k = I(x_k) > 0.$$

注意到 a 是以 $2T > 0$ 为周期的, 而 $x_k(t)$ 是以 $2kT$ 为周期的, $x_k(t-2jT)|_{[-kT,kT]}$ ($j \in \mathbb{Z}$) 都是方程（16.10）的具有相同周期、相同临界值的解. 从中我们可以选取这样一个 x_k 满足:

$$\max_{t \in [-T,T]} x_k(t) = \max_{t \in [-kT,kT]} x_k(t). \tag{16.12}$$

分如下几步来证明.

1° $\exists\, M > 0$, $c_k \leqslant M$, $\forall\, k \in \mathbb{N}$.

事实上取 $\varphi_1 \in X_1$, 满足 $\varphi_1(\pm T) = 0$, 以及 $I_1(\varphi_1) \leqslant 0$. 令

$$M = \max_{t \in [0,1]} I_1(t\varphi_1)$$

以及

$$z_k(t) = \begin{cases} \varphi_1(t), & t \in [-T, T], \\ 0, & t \in [-kT, kT] \backslash [-T, T], \end{cases}$$

则有

$$z_k \in X_k, \quad I_k(z_k) = I_1(\varphi_1) \leqslant 0.$$

于是得到估计

$$c_k \leqslant \max_{t \in [0,1]} I_k(tz_k) = \max_{t \in [0,1]} I_1(t\varphi_1) = M.$$

2° $\exists\, M_1 > 0$ 使得 $\|x_k\|_k \leqslant M_1$.

在 (16.11) 中取 $\varphi = x_k$, 得到

$$\int_{-kT}^{kT} \left[\dot{x}_k^2 + x_k^2 - \mu a(t)|x_k(t)|^\mu\right] dt = 0,$$

与等式

$$c_k = I_k(x_k) = \frac{1}{2}\int_{-kT}^{kT}(\dot{x}_k^2 + x_k^2)dt - \int_{-kT}^{kT} a(t)|x_k(t)|^\mu dt$$

联立得到

$$\left(\frac{\mu}{2} - 1\right)\int_{-kT}^{kT} a(t)|x_k(t)|^\mu dt = c_k,$$

于是有

$$\|x_k\|_k^2 = 2\left(I_k(z_k) + \int_{-kT}^{kT} a(t)|x_k(t)|^\mu dt\right) \leqslant \left(2 + \frac{2}{\mu - 2}\right) c_k \leqslant M_1.$$

3° 我们要证明存在解 $x \in H^1(\mathbb{R}^1)$. 注意到对于任意 $x \in H^1_{\text{loc}}$, 有不等式

$$|x(t)| \leqslant |x(s)| + \int_s^t |\dot{x}(r)|dr,$$

从而

$$|x(t)| \leqslant \int_{t-\frac{1}{2}}^{t+\frac{1}{2}}\left(|x(s)| + \int_s^t |\dot{x}(r)|dr\right)ds \leqslant 2\int_{t-\frac{1}{2}}^{t+\frac{1}{2}}(\dot{x}(r)^2 + x(r)^2)dr,$$

所以存在与 k 无关的常数 C, M_2 使得

$$\|x\|_{L^\infty[-kT,kT]} \leqslant C\|x\|_k^2 \leqslant M_2.$$

代回方程 (16.11) 就得到与 k 无关的常数 M_3 使得

$$\|x_k\|_{C^2[-kT,kT]} \leqslant M_3.$$

这表明在任意有界区间 $(-R, R)$ 上, x_k 连同它们的导数一致收敛到一个连续可微函数 y, 即:

$$x_k \to y, \text{ 于 } C^1[-R, R].$$

于是对于任意 $k \in \mathbb{N}$,

$$\int_{-kT}^{kT}[\dot{y}^2 + y^2]dt \leqslant M_1,$$

因为 k 是任意的, 所以

$$\int_{-\infty}^{\infty}[\dot{y}^2 + y^2]dt \leqslant M_1, \tag{16.13}$$

就得到 $y \in H^1(\mathbb{R}^1)$, 并且它在全直线上满足方程 (16.11) 以及 $y(0) = 0$.

4° 证明求得的解 y 是同宿轨. 即需证明: $y(t) \to 0$, 当 $t \to \pm\infty$. 我们采用 y 的 Fourier 变换:

$$\tilde{y}(\xi) = \frac{1}{2\pi}\int_{-\infty}^{\infty} y(t)e^{i\xi t}dt.$$

由 Plancherel 定理

$$\int_{-\infty}^{\infty}(1+|\xi|^2)|\tilde{y}(\xi)|^2 d\xi = \int_{-\infty}^{\infty}[\dot{y}^2 + y^2]dt \tag{16.14}$$

与 Schwarz 不等式得到

$$\int_{-\infty}^{\infty}|\tilde{y}(\xi)|d\xi \leqslant \left(\int_{-\infty}^{\infty}(1+|\xi|^2)|\tilde{y}(\xi)|^2 d\xi\right)^{1/2}\left(\int_{-\infty}^{\infty}(1+|\xi|^2)^{-1}d\xi\right)^{1/2} < \infty,$$

再由 Reimann-Lebesgue 定理即得

$$y(t) \to 0 \quad (t \to \pm\infty).$$

5° 最后我们来解释为什么这个解 y 是非平凡的. 主要的想法是要证明存在一个常数 $\delta > 0$, 使得

$$\|x_k\|_{L^\infty[-kT,kt]} \geqslant \delta. \tag{16.15}$$

为此定义函数

$$\phi(r) = \max_{t\in[-T,T],\,|x|\leqslant r} a(t)|x|^{\mu-2}.$$

这是一个单调不减的连续函数, 满足

$$\begin{cases} \phi(r) > 0, & r > 0, \\ \phi(r) \to \infty, & r \to \infty. \end{cases}$$

因为 x_k 是 (16.11) 的解, 所以

$$\|x_k\|_k^2 = \mu\int_{-kT}^{kT} a(t)|x_k(t)|^\mu dt.$$

ϕ 与 $\|x_k\|_k$ 的关系是

$$\|x_k\|_k^2 \leqslant \mu\phi(\|x_k\|_{L^\infty[-kT,kT]})\int_{-kT}^{kT}|x_k(t)|^2 dt \leqslant \mu\phi(\|x_k\|_{L^\infty[-kT,kT]})\|x_k\|_k^2,$$

于是得到

$$\phi(\|x_k\|_{L^\infty[-kT,kT]}) \geqslant \frac{1}{\mu}.$$

适当取 δ 即得 (16.15).

因为有 (16.12), 所以

$$\max_{t\in[-T,T]} |x_k(t)| \geqslant \delta,$$

便有

$$\max_{t\in\mathbb{R}^1} |y(t)| \geqslant \delta.$$

定理 16.4 设 $\mu > 2, a \in C(\mathbb{R}^1)$ 是一个 $2T$-周期函数 $(T > 0)$, 又设 $\exists\alpha > 0$, 使得 $a(t) \geqslant \alpha, \forall t \in \mathbb{R}^1$, 则方程 (16.11) 具有非平凡的同宿轨.

第十七讲　测地线与极小曲面

§17.1　测地线

设 (M, g) 是一个 Riemann 流形,

$$\gamma: \ [0,1] \longrightarrow M, \quad \dot{\gamma} \neq \theta$$

称为测地线, 如果加速度向量场

$$\frac{D}{dt}\frac{d\gamma}{dt} = 0,$$

即, 沿曲线 γ, 其速度向量场是平行的. 从而沿 γ,

$$\frac{d}{dt} g\left(\frac{d\gamma}{dt}, \frac{d\gamma}{dt}\right) = 2g\left(\frac{D}{dt}\left(\frac{d\gamma}{dt}\right), \frac{d\gamma}{dt}\right) = 0 \Longleftrightarrow \left\|\frac{d\gamma}{dt}\right\| = g(\dot{\gamma}, \dot{\gamma})^{1/2} = \text{const}.$$

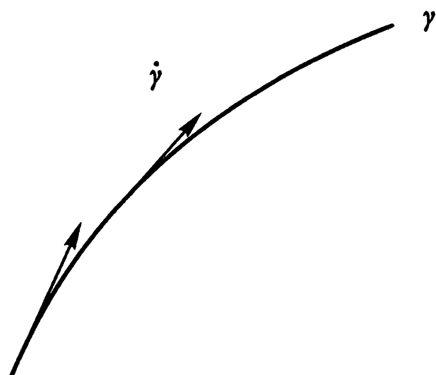


图 17.1

在局部坐标下: $\gamma=(u^1,\cdots,u^n)$, 测地线方程为

$$\frac{d^2u^k}{dt^2}+\sum_{i,j=1}^{n}\Gamma_{ij}^k(u)\frac{du^i}{dt}\cdot\frac{du^j}{dt}=0,\quad 1\leqslant k\leqslant n,$$

其中 Γ_{ij}^k 是 M 上的 Christoffel 符号,

$$\Gamma_{ij}^k=\frac{1}{2}\sum_{l=1}^{n}g^{kl}\left(\frac{\partial g_{il}}{\partial u^j}+\frac{\partial g_{jl}}{\partial u^i}-\frac{\partial g_{ij}}{\partial u^l}\right),$$

$$g_{ij}=g\left(\frac{\partial}{\partial u^i},\frac{\partial}{\partial u^j}\right),\quad 1\leqslant i,j,k\leqslant n.$$

$$\sum_{j=1}^{n}g^{ij}g_{jk}=\delta_k^i,$$

- **弧长泛函与能量泛函**

对于 (M,g) 上的任意一条 C^1 曲线 γ, 或更一般地, $\gamma\in W^{1,1}([0,1],M)$,

$$L(\gamma)=\int_0^1\left\|\frac{d\gamma}{dt}(t)\right\|dt$$

称为弧长泛函. 当 $\gamma\in H^1([0,1],M)$ 时, 称

$$E(\gamma)=\frac{1}{2}\int_0^1\left\|\frac{d\gamma}{dt}(t)\right\|^2dt$$

为能量泛函.

由 Schwarz 不等式得

$$L(\gamma)^2\leqslant 2E(\gamma),$$

而且

$$\text{“}=\text{”}\Longleftrightarrow\left\|\frac{d\gamma}{dt}(t)\right\|=\text{const}.$$

- **L 在参数微分同胚意义下是不变的**

$$\begin{aligned}L(\gamma\circ s)&=\int_0^T\left\|\frac{d\gamma\circ s}{d\tau}(\tau)\right\|d\tau\\&=\int_0^T\left\|\frac{d\gamma}{dt}(s(\tau))\right\|\left\|\frac{ds}{d\tau}(\tau)\right\|d\tau\\&=\int_0^1\left\|\frac{d\gamma}{dt}(t)\right\|dt\\&=L(\gamma).\end{aligned}$$

但 E 不是微分同胚不变的!

- 规范弧长函数

$$l(t)=\frac{1}{L(\gamma)}\int_0^t\left\|\frac{d\gamma}{ds}(s)\right\|ds,\quad t\in[0,1].$$

$l:\ [0,1]\longrightarrow[0,1]$ 是微分同胚, 令 $s=l^{-1}$, 则 $\gamma\circ s:\ [0,1]\longrightarrow M$ 满足

$$\left\|\frac{d(\gamma\circ s)}{dl}(l)\right\|=\left\|\frac{d\gamma}{dt}\circ s(l)\right\|\frac{ds}{dl}(l)=\left\|\frac{d\gamma}{dt}(t)\right\|\left(\frac{dl}{dt}(t)\right)^{-1}=L(\gamma).$$

若曲线 γ 以规范弧长为参数, 即 $\gamma\circ s$ 为其参数化, 则

$$L(\gamma)^2=2E(\gamma).$$

给定 $P_0,P_1\in M$, 令

$$\Gamma=\left\{\gamma\in H^1([0,1],M)\mid\gamma(i)=P_i,\ i=0,1\right\}.$$

因此我们有

$$\boxed{\inf_{\gamma\in\Gamma}L(\gamma)=\inf_{\gamma\in\Gamma}\sqrt{2E(\gamma)}.}$$

寻求连接两定点 P_0 与 P_1 长度最短的曲线, 等价于寻求连接 P_0 与 P_1 能量最小的曲线.

因此, 长度最短的曲线 γ 极小化能量.

于是, γ 满足 E 的 E-L 方程

$$\frac{d}{dt}L_P\left(t,u(t),\dot u(t)\right)=L_u\left(t,u(t),\dot u(t)\right),$$

其中 $L(t,u,p)=\frac{1}{2}\sum_{i,j=1}^n g_{ij}(u)p_ip_j$. 即

$$\begin{aligned}
&\sum_{j=1}^n\left[\frac{d}{dt}\left(g_{ij}(u(t))\dot u^j(t)\right)-\frac{1}{2}\sum_{k=1}^n\frac{\partial g_{kj}}{\partial u^i}\left(u(t)\right)\dot u^k(t)\dot u^j(t)\right]=0,\\
\Longleftrightarrow&\sum_{j=1}^n\left[2g_{ij}\ddot u^j+2\sum_{k=1}^n\frac{\partial g_{ij}}{\partial u^k}\dot u^k\dot u^j-\sum_{k=1}^n\frac{\partial g_{kj}}{\partial u^i}\dot u^k\dot u^j\right]=0,\\
\Longleftrightarrow&\ddot u^i+\frac{1}{2}\sum_{l=1}^n\sum_{k=1}^n\sum_{j=1}^n g^{il}(2g_{lj,k}\dot u^k\dot u^j-g_{kj,l}\dot u^k\dot u^j)=0,\\
\Longleftrightarrow&\ddot u^i+\frac{1}{2}\sum_{l=1}^n\sum_{k=1}^n\sum_{j=1}^n g^{il}(g_{lj,k}+g_{kl,j}-g_{jk,l})\dot u^k\dot u^j=0,\\
\Longleftrightarrow&\ddot u^i+\sum_{k,j=1}^n\Gamma_{jk}^i\dot u^k\dot u^j=0,\quad i=1,\cdots,n.
\end{aligned}$$

这正是测地线方程. 所以长度最短的曲线是测地线!

● **测地线的存在性**

设 (M, g) 是一个紧 Riemann 流形, 给定 $P_0, P_1 \in M$. 问: 是否存在连接 P_0, P_1 的测地线?

按 Nash 嵌入定理, 将 (M, g) 等距地嵌入到 $\mathbb{R}^N$(N 足够大). 对于泛函

$$E(u) = \frac{1}{2}\int_0^1 \|\dot{u}(t)\|_{\mathbb{R}^N}^2 dt,$$

考虑集合

$$\begin{aligned} S &= \{u \in H^1([0,1], \mathbb{R}^N) \mid u(0) = P_0, u(1) = P_1, u(t) \in M, \forall t \in [0,1]\} \\ &= C_*([0,T], M) \cap H^1([0,1], \mathbb{R}^N), \end{aligned}$$

其中

$$C_*([0,T], M) = \{u \in C([0,T], M) \mid u(i) = P_i,\ i = 0, 1\}.$$

这是因为嵌入映射 $i:\ H^1 \to C$ 是连续的.

更强地, 我们还有

$$\begin{aligned} \|u(t_1) - u(t_2)\|_{\mathbb{R}^N} &= \left\|\int_{t_1}^{t_2} \dot{u}(t)dt\right\| \\ &\leqslant \int_{t_1}^{t_2} \|\dot{u}(t)\|dt \\ &\leqslant |t_1 - t_2|^{1/2}\left(\int_{t_1}^{t_2} \|\dot{u}(t)\|^2 dt\right)^{1/2} \\ &\leqslant (2E(u))^{1/2}|t_1 - t_2|^{1/2}, \end{aligned}$$

这表明 S 是 $H^1([0,T], \mathbb{R}^N)$ 中的弱闭集.

由直接方法, E 有极小点, 它也必是 L 的极小点.

我们证明了如下定理.

定理 17.1 设 (M, g) 是一个紧 Riemann 流形, 又设 $P_0, P_1 \in M$, 则存在一条连接 P_0 与 P_1 的长度最短的曲线, 这曲线满足测地线方程. □

注 17.1 我们不去直接寻求 $L(\gamma)$ 的极小点, 是因为 $L(\gamma)$ 对 $[0,1]$ 上的微分同胚不变, 而这微分同胚群很大, 没有紧性.

相反, $E(\gamma)$ 不是微分同胚不变的, 有紧性可以利用. 在 Hilbert 的原始证明中, 采用 “弧长为曲线参数” 也是为了避开这个微分同胚群, 本质上和对能量求极小值是一样的. □

§17.2 极小曲面

几何上把平均曲率为零的曲面定义为极小曲面.

对于区域 $\Omega \subset \mathbb{R}^2$ 上的曲面, 我们采用参数方程:

$$X: \Omega \to \mathbb{R}^n, \quad (x,y) \mapsto X = (X^1, \cdots, X^n).$$

在等温坐标下, 极小曲面的参数表示 X 满足

$$\begin{cases} \Delta X^i = 0, \quad 1 \leqslant i \leqslant n & \text{(调和方程)}, \\ \sum\limits_{i=1}^{n} [(X_x^i)^2 - (X_y^i)^2] = \sum\limits_{i=1}^{n} X_x^i \cdot X_y^i = 0 & \text{(弱保形条件)}. \end{cases} \tag{17.1}$$

例 17.1 悬链旋转面

$$\begin{cases} x = \operatorname{ch} u \sin v, \\ y = \operatorname{ch} u \cos v, \\ z = u. \end{cases}$$

其名称之由来是与如下问题有关的.

- **极小曲面的 Plateau 问题.**

在 $\mathbb{R}^n$ 中给定一条 Jordan 曲线 Γ (见图 17.2). 求一 (盘状) 曲面 $X: \bar{D} \to \mathbb{R}^n$ 张在 Γ 上使其面积最小. 记

$$\bar{D} = \{w = (x,y) \in \mathbb{R}^2 \mid |w| \leqslant 1\},$$
$$X = (X^1(w), \cdots, X^n(w)).$$

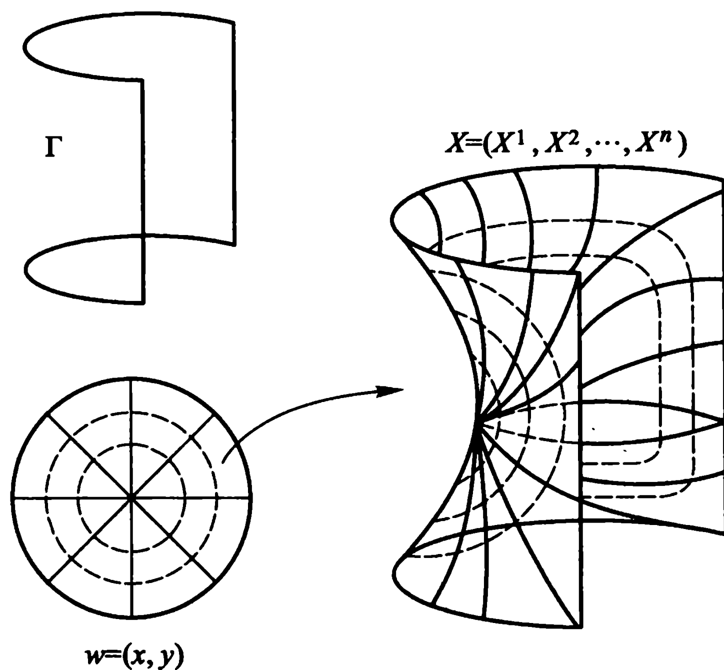


图 17.2

注意到两个向量决定的面积平方为

$$|X_x \wedge X_y|^2 = |X_x|^2|X_y|^2 - (X_x \cdot X_y)^2,$$

所以曲面 X 的面积是

$$A(x) = \int_D |X_x \wedge X_y| dx \wedge dy.$$

• **边值条件**. 令 $\hat{X} = X|_{\partial D}$. 我们假设 $\hat{X} : \partial D \to \Gamma$ 是一个定向参数化, 即保持定向的同胚.

记

$$C(\Gamma) = \{X \in H^1(\bar{D}, \mathbb{R}^n) | \hat{X} = X|_{\partial D} \in C(\partial D, \Gamma) \text{ 是弱单调的参数化}\}.$$

在从变分角度来研究极小曲面之前, 我们首先要对 Γ 添加条件, 以使

$$\inf_{X \in C(\Gamma)} A(X) < \infty,$$

才能保证 $C(\Gamma) \neq \varnothing$.

例如设 Γ 是可求长的, 然后再证明存在 $X_0 \in C(\Gamma)$, 使得

$$A(X_0) = \inf_{X \in C(\Gamma)} A(X).$$

• **A 是微分同胚不变的**.

令 $z = (u, v)$, 以及 $z = z(w)$, 则

$$\frac{dz}{dw} = \begin{pmatrix} \partial_x u & \partial_y u \\ \partial_x v & \partial_y v \end{pmatrix}.$$

因此,

$$X_x \wedge X_y \; dx \wedge dy = X_u \wedge X_v \; du \wedge dv.$$

微分同胚群 $\mathrm{Diff}(\bar{D}, \bar{D})$ 太大, 当然无紧性可言! 因此用直接方法对面积泛函 A 求极小有困难. 仿照测地线问题, 我们转向 Dirichlet 积分

$$D(X) = \frac{1}{2} \int_D |\nabla X|^2 dx \wedge dy,$$

其中 $\nabla X = (X_x^i, X_y^i)$, $1 \leqslant i \leqslant n$.

• **保形变换**. 一个微分同胚

$$g:\ \mathbb{R}^2 \to \mathbb{R}^2, \quad w \mapsto z$$

称为是保形的, 如果它满足如下弱保形条件:

$$\boxed{|g_x|^2 - |g_y|^2 = g_x \cdot g_y = 0.}$$

直接计算可知: Dirichlet 积分虽然不是微分同胚不变的, 但却是保形不变的!

$$\int_D |\nabla_w X|^2 dx \wedge dy = \int_D |\nabla_z X|^2 du \wedge dv.$$

注意到下列的简单关系:

$$|\alpha \wedge \beta| \leqslant \frac{1}{2}(|\alpha|^2 + |\beta|^2), \quad \text{“=”} \Longleftrightarrow |\alpha|^2 - |\beta|^2 = \alpha \cdot \beta = 0;$$
$$A(X) \leqslant D(X), \quad \text{“=”} \Longleftrightarrow X \text{ 是弱保形的}.$$

我们可以得到如下重要的结论.

定理 17.2 (Morrey-Lichtenstein) 设 $\Gamma \in C^2$ 是一条 Jordan 曲线, 则 $C(\Gamma) \neq \varnothing$, 并且

$$\boxed{\inf_{X \in C(\Gamma)} A(X) = \inf_{X \in C(\Gamma)} D(X).}$$

为了直接进入存在性的证明, 我们把这个结论的证明留到本讲的最后.

把极小化面积泛函转化到极小化 Dirichlet 积分的好处是: 保形变换群比微分同胚群小得多.

保形变换群. 一切保形变换构成一个群, 它由三个实参数生成:

$$G = \left\{ g(w) = e^{i\phi} \frac{w + a}{1 + \bar{a}w} \,\middle|\, |a| < 1, \phi \in \mathbb{R}^1 \right\}.$$

注意到当 $a = a_j \to 1$ 时, $g_j(w)$ 集中到单位圆周上的一点. 这表明 $\forall X \in C(\Gamma)$, 轨道集 $\{X \circ g \,|\, g \in G\}$ 在 $H^1(D, \mathbb{R}^n)$ 的弱拓扑下的闭包含有常值函数. 后者不可能是 Γ 的弱单调参数化, 所以 $C(\Gamma)$ 就不可能在 H^1 中成为弱闭集. 即使我们把泛函换到了 Dirichlet 积分, 仍然没能绕开 $C(\Gamma)$ 的弱闭性问题.

我们采用的办法是进一步消除保形群 G 不变的影响: 在 $C(\Gamma)$ 上添加限制, 以 “商” 去 G.

在 Γ 上任意取定定向的三点 P_0, P_1, P_2, 令

$$C^*(\Gamma) = \left\{ X \in C(\Gamma) \,\middle|\, X\left(e^{\frac{2k\pi}{3} i}\right) = P_k, 0 \leqslant k \leqslant 2 \right\}.$$

我们把 $X\left(e^{\frac{2k\pi}{3}i}\right)=P_k, 0\leqslant k\leqslant 2$ 称为三点条件. 由于

$$\boxed{\inf_{X\in C(\Gamma)} A(X)=\inf_{X\in C(\Gamma)} D(X)=\inf_{X\in C^*(\Gamma)} D(X),}$$

我们便转求

$$\boxed{\min\{D(X),\ X\in C^*(\Gamma)\}.}$$

如果 X 是极小化函数, 那么它应满足:

(1)

$$\frac{d}{d\varepsilon}D(X+\varepsilon\varphi)\bigg|_{\varepsilon=0}=0,$$

(2)

$$\frac{d}{d\varepsilon}D(X\circ g_\varepsilon^{-1},g_\varepsilon(D))\bigg|_{\varepsilon=0}=\frac{d}{d\varepsilon}\frac{1}{2}\int_{g_\varepsilon(D)}\left|\nabla_z(X\circ g_\varepsilon^{-1}(z))\right|^2du\wedge dv|_{\varepsilon=0}=0,$$

其中 g_ε: $D\to g_\varepsilon(D)$ 是微分同胚.

理由如下. 虽然 $C^*(\Gamma)$ 不是微分流形, 但 $\forall X\in C^*(\Gamma),\forall\varphi\in C_0^\infty(D,\mathbb{R}^n), X+\varphi\in C^*(\Gamma)$, 所以 (1) 成立. 由此推出 E-L 方程

$$\Delta X=0\quad 于\ D,$$

这是 (17.1) 的第一个方程. □

至于 (2), 我们用反证法. 如果 $\frac{d}{d\varepsilon}D(X\circ g_\varepsilon^{-1},g_\varepsilon(D))|_{\varepsilon=0}\neq 0$, 那么, 必存在一个微分同胚 $\bar{g}_\varepsilon\in C^1(\bar{D},\mathbb{R}^2)$, 使得 $\bar{X}=X\circ\bar{g}_\varepsilon$ 满足

$$D(\bar{X}_\varepsilon,\bar{g}_\varepsilon(D))<D(X).$$

但 $\bar{g}_\varepsilon(D)$ 是单连通的, 由 Riemann 定理, 存在保形映射. $h_\varepsilon: D\to\bar{g}_\varepsilon(D)$. 令 $\tilde{X}_\varepsilon=\bar{X}_\varepsilon\circ h_\varepsilon$, 则 $\tilde{X}_\varepsilon\in C(\Gamma)$. 再由 Dirichlet 积分的保形不变性,

$$D(\tilde{X}_\varepsilon)=D(\bar{X}_\varepsilon,\bar{g}_\varepsilon(D))<D(X).$$

这是一个矛盾!

由 (2), 再利用第 8 讲例 8.4 的类似推理, 得出在 D 上,

$$\begin{cases}|X_x|^2=|X_y|^2,\\ X_x\cdot X_y=0,\end{cases}$$

这是 (17.1) 的第二个方程.

事实上, 设 $\frac{dg_\varepsilon}{d\varepsilon}|_{\varepsilon=0}=\tau$, $\tau=(\tau^1,\tau^2)$. 应用第 8 讲的 Noether 公式于 $L(p)=\frac{1}{2}\left(\sum_{i=1}^n(p_1^i)^2+(p_2^i)^2\right)$, 我们有

$$\int_D \mathrm{div}(L\tau+L_{p^i}\cdot\varphi^i)dxdy=0,$$

其中 $\varphi^i=-(X_x^i\tau^1+X_y^i\tau^2)$.

如今

$$\begin{aligned}&L\tau+L_{p^i}\cdot\varphi^i\\&=\frac{1}{2}(|X_x|^2+|X_y|^2)(\tau^1,\tau^2)-\sum_{i=1}^n(X_x^i\tau^1+X_y^i\tau^2)(X_x^i,X_y^i)\\&=\left(-\frac{1}{2}(|X_x|^2-|X_y|^2)\tau^1-X_x\cdot X_y\tau^2,\ -X_x\cdot X_y\tau^1+\frac{1}{2}(|X_x|^2-|X_y|^2)\tau^2\right),\end{aligned}$$

所以

$$\mathrm{div}(L\tau+L_{p^i}\cdot\varphi^i)=\frac{1}{2}(|X_x|^2-|X_y|^2)(\tau_y^2-\tau_x^1)-X_x\cdot X_y(\tau_y^1+\tau_x^2).$$

我们证明了

$$\begin{aligned}&\frac{d}{d\varepsilon}D(X\circ g_\varepsilon^{-1},g_\varepsilon(D))|_{\varepsilon=0}\\&=\frac{d}{d\varepsilon}\frac{1}{2}\int_{g_\varepsilon(D)}\left|\nabla_z(X\circ g_\varepsilon^{-1})\right|^2du\wedge dv\\&=-\frac{1}{2}\int_D[(|X_x|^2-|X_y|^2)(\tau_x^1-\tau_y^2)+2X_x\cdot X_y(\tau_y^1+\tau_x^2)]dxdy.\end{aligned}$$

因为 $\lambda=\tau_x^1-\tau_y^2$ 与 $\mu=\tau_y^1+\tau_x^2$ 可以是任意的, 所以

$$|X_x|^2-|X_y|^2=X_x\cdot X_y=0\quad \text{a.e. 于 } D.$$ □

我们证明了: 能量泛函的极小化函数是极小曲面方程的广义解.

• 当把 $\mathbb{R}^n$ 换成一般的 Riemann 流形 (N^n,h) 时, Γ 是嵌入在 N 中的一条 Jordan 曲线. 相应的方程变为

$$\begin{cases}\mathrm{trace}(\nabla dX)=0,\\ h(X_x,X_x)-h(X_y,X_y)=h(X_x,X_y)=0.\end{cases}$$

作等距嵌入 $N\hookrightarrow\mathbb{R}^N$, 还可以把这方程变为

$$\begin{cases}\Delta X=A(X)(\nabla X,\nabla X),\\ |X_x|^2-|X_y|^2=X_x\cdot X_y=0,\end{cases}$$

其中 $A(X)(\cdot,\cdot)$ 是曲面 X 的第二基本形式.

极小曲面 Plateau 问题解的存在性的证明是 Douglas (1931), Rado (1933), Courant (1945), M. Struwe (1988) 等给出的. 相应的正则性则是由 S. Hilbrandt (1969, 1971) 证明的: 若 $\Gamma \in C^{2,\alpha}$, 则 $X \in C^{2,\alpha}(\overline{\Omega}, \mathbb{R}^N), 0 < \alpha < 1$.

存在性的证明. 特别值得注意的是, 为了极小化 $D(X)$, 我们并不需要考虑整个函数 X, 而只需要考虑它的边值 $\hat{X}$. 这是因为从 E-L 方程知道, 调和函数实现 Dirichlet 积分 $D(X)$ 的极小值, 而调和函数又完全由它的边值决定. 事实上, 如果我们把 $\hat{X}$ 用 Fourier 级数展开:

$$\hat{X}(e^{i\theta}) = a_0 + \sum_{k=1}^{\infty}(a_k \cos k\theta + b_k \sin k\theta), \tag{17.2}$$

其中

$$a_0 = \frac{1}{2\pi}\int_0^{2\pi} \hat{X}(e^{i\theta})d\theta, \quad a_k = \frac{1}{\pi}\int_0^{2\pi} \hat{X}(e^{i\theta}) \cos k\theta d\theta,$$

$$b_k = \frac{1}{\pi}\int_0^{2\pi} \hat{X}(e^{i\theta}) \sin k\theta d\theta, \quad k = 1, 2, \cdots,$$

$$\sum_{k=1}^{\infty} k(|a_k|^2 + |b_k|^2) < \infty,$$

那么

$$X(re^{i\theta}) = \frac{1}{2}a_0 + \sum_{k=1}^{\infty}(a_k \cos k\theta + b_k \sin k\theta)r^k. \tag{17.3}$$

由此可见, 我们只需把

$$D(X) = \pi \sum_{k=1}^{\infty} k(|a_k|^2 + |b_k|^2)$$

看成集合 $C^*(\Gamma) \subset X = \{\hat{X} = X|_{\partial D} \mid D(X) < \infty\}$ 上的泛函, 并在空间 X 上定义范数

$$\|\hat{X}\| = \left(\frac{1}{4}a_0^2 + \sum_{k=1}^{\infty} k(|a_k|^2 + |b_k|^2)\right)^{1/2},$$

使它成为一个 Hilbert 空间. D 显然是弱序列下半连续的. 注意到 $\hat{X}$ 的值域落在 Γ 上, 这表明 $|a_0|$ 是有界的, 从而 $D(X)$ 是强制的.

为了使用直接方法, "紧性" 还是最需要关注的.

如今 Γ 是给定的, $\hat{X}$ 只是 Γ 的一个参数化.

对于一族极小化函数序列来说, 它们的连续模量其实都是 Γ 经过同胚变换后的连续模量. 为了得到极小化函数序列的等度连续性, 关键之处在于证明: 对于能量有界的这样一族参数化, 不能把区间集中到一个点去. 下列引理起着重要的作用.

引理 17.1 (Courant-Lebesgue) 设 $X \in H^1(D, \mathbb{R}^n)$, 则 $\forall w \in \bar{D}, \forall \delta \in (0,1), \exists \rho \in [\delta, \sqrt{\delta}]$, 使得

$$\int_{C'_\rho} |\partial_s X|^2 ds \leqslant 4\frac{D(X)}{\rho|\ln \delta|},$$

其中 ∂_s 是切向导数, 而 C'_ρ 是圆心在单位圆周上以 ρ 为半径的圆弧 (图 17.3).

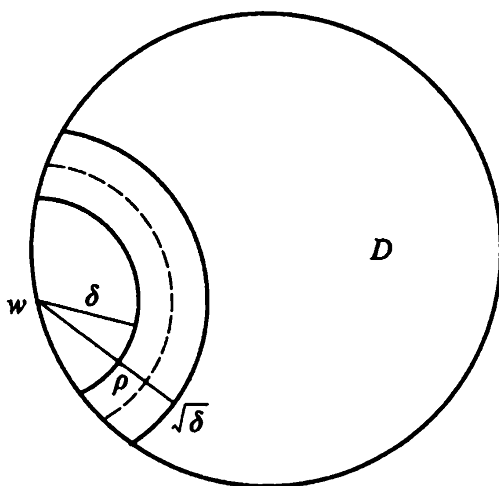


图 17.3

证明 由 Fubini 定理, 对几乎所有的 $\rho \in (0,1), |\partial_s X| \in L^2(C_\rho)$, 而且

$$\begin{aligned} 2D(X) &\geqslant \int_{B_{\sqrt{\delta}}(w)\setminus B_\delta(w)} |\nabla X|^2 dz \\ &\geqslant \int_\delta^{\sqrt{\delta}} \int_{C_\rho} |\partial_s X|^2 ds d\rho \\ &\geqslant \operatorname*{ess\,inf}_{\delta<\rho<\sqrt{\delta}} \rho \int_{C_\rho} |\partial_s X|^2 ds \int_\delta^{\sqrt{\delta}} \frac{d\rho}{\rho}. \end{aligned}$$

再由

$$\int_\delta^{\sqrt{\delta}} \frac{d\rho}{\rho} = \frac{1}{2}|\ln \delta|,$$

即得结论. □

我们利用它来推出如下引理.

引理 17.2 $\forall M > 0$, 集合

$$C_M = \left\{\hat{X}:\ \partial D \to \Gamma \middle| \hat{X} = X|_{\partial D},\ X \in C^*(\Gamma), D(X) \leqslant M\right\}$$

是等度连续的.

证明 因为 Γ 是 Jordan 曲线, 所以 $\forall \varepsilon > 0, \exists d > 0, \forall P \neq P'$,

$$\boxed{|P - P'| < d} \Rightarrow \boxed{\Gamma\backslash\{P, P'\}\text{ 中, 至少有一个分量 diam} < \varepsilon.}$$

1° $\forall \varepsilon > 0$, 设 $\varepsilon < \min\limits_{i\neq j}|P_i - P_j|$. 取 $d > 0$ 如上, 取 $\delta > 0$ 使 $8\pi M/|\ln\delta| < d^2$. $\forall X \in C_M, \forall w \in \partial D$, 取 $C_\rho = \partial D \cap \partial B_\rho(w)$, 由 Courant-Lebesgue 引理,

$$\begin{aligned}|X(w_1) - X(w_2)|^2 &\leqslant [\text{diam}X(C_\rho)]^2 \\ &\leqslant \left(\int_{C_\rho} |\partial_s X| ds\right)^2 \\ &\leqslant L(C_\rho)\int_{C_\rho} |\partial_s X|^2 ds \\ &\leqslant \frac{8\pi M}{|\ln\delta|},\ \forall w_1, w_2 \in C_\rho.\end{aligned}$$

因此, $\{w_1, w_2\} = \partial D \cap \partial B_\rho(w) \Longrightarrow$ 在 $\Gamma\backslash\{X(w_1), X(w_2)\}$ 中, 有一个分量, 其 diam $< \varepsilon$.

2° 取 $\delta > 0$, 使得 $\forall z \in \partial D$, 至少有两个 k, 满足

$$\left|z - e^{\frac{2\pi k}{3}i}\right| \geqslant \sqrt{\delta}.$$

设 $\partial D\backslash\{w_1, w_2\} = C_1 \cup C_2$, $C_1 \cap C_2 = \{w_1, w_2\}$, 则 C_1, C_2 中至少有一个, 例如说 C_1, 至多仅含 $\left\{e^{\frac{2k\pi}{3}i}\middle| k = 1, 2, 3\right\}$ 中的一个. 于是,

$$\text{diam}\hat{X}(C_1) < \varepsilon.$$

$\forall z, z' \in \partial D$, 当 $|z - z'| < \delta$, 可以取 $w_0 \in \partial D, \rho > 0$, 使 $z, z' \in C_1$. 从而

$$|\hat{X}(z) - \hat{X}(z')| \leqslant \text{diam}\hat{X}(C_1) < \varepsilon \Longrightarrow C_M\text{ 是等度连续的.}$$ □

定理 17.3 设 $\Gamma \in C^2$ 是一条 Jordan 曲线, 则 $\exists X_0 \in C^*(\Gamma)$, 使得

$$D(X_0) = \inf_{X\in C^*(\Gamma)} D(X).$$

证明 我们在 $C^*(\Gamma)$ 上考察泛函 $D(X)$,

$$\hat{X} \leftrightarrow X \mapsto D(X),$$

其中 $\hat{X}$ 与 X 的关系由 (17.2) 与 (17.3) 决定. 我们已经知道它是弱序列下半连续的、强制的, 剩下只要证明 $C^*(\Gamma)$ 是弱闭的就够了. 设 $\{\hat{X}_j\}$ 弱收敛到 $\hat{X}_0$, 则模有界. 由引理 17.2, 这序列等度连续. 又因为 $\hat{X}$ 的值域在 Γ 上, 所以函数 $\{\hat{X}_j\}$ 是一致有界的, 从而有子列 $\{\hat{X}_{j'}\}$ 一致收敛到一个连续函数, 这表明函数 $\hat{X}_0 : \partial D \to \Gamma$ 是连续的, 而且还是弱单调的. 此外满足三点条件, 即 $X_0 \in C^*(\Gamma)$.

因此极小曲面 Plateau 问题的解在 $C(\Gamma)$ 上存在. 由 Weyl 定理, 这解在 D 内是 C^∞ 的. 至于在 $\bar{D}$ 上的正则性则已超出本讲义范围. □

现在我们回过来证明定理 17.2, 即 $A(X)$ 与 $D(X)$ 有共同的下确界. 我们已经知道 $A(X) \leqslant D(x)$, 而且 "=" $\Leftrightarrow$ X 是弱保形的.

引理 17.3 设 $\Gamma \in C^2$ 是一条 Jordan 曲线, 而且 $X \in C(\Gamma) \cap C^2(\bar{D}, \mathbb{R}^n)$, 则 $\forall \varepsilon > 0$, $\exists g \in H^1 \cap C(\bar{D}, \mathbb{R}^2)$, $g : \bar{D} \to \bar{D}$ 是满射, 并且 $g : \partial D \to \partial D$ 单调, 使得 $X_\varepsilon \circ g$是弱保形的, 其中 $X_\varepsilon(x, y) = (X(x, y), \varepsilon x, \varepsilon y) \in C^2(\bar{D}, \mathbb{R}^{n+2})$.

证明 令

$$S = \{g \in C^1(\bar{D}, \mathbb{R}^2) \mid g : \bar{D} \to \bar{D} \text{ 是微分同胚}, g(e^{\frac{2k\pi}{3}i}) = e^{\frac{2k\pi}{3}i}, 1 \leqslant k \leqslant 3\},$$

$\bar{S}$ 为 S 在 $H^1(D, R^2)$ 中的弱闭包. 再令

$$E(g) = D(X_\varepsilon \circ g) = \frac{1}{2}\int_D |(\nabla X_\varepsilon) \circ g \cdot \nabla g|^2 dxdy.$$

由于

$$E(g) \geqslant \varepsilon D(g),$$

所以 E 是 $\bar{S}$ 上弱序列下半连续、强制的泛函. 按定理 17.3 的证明, E 在 $\bar{S}$ 上有极小点 g_0. 即

$$D(X_\varepsilon \circ g_0) = E(g_0) \leqslant D(X_\varepsilon \circ g), \quad \forall g \in \bar{S}.$$

再利用 (2) 导出 $X_\varepsilon \circ g_0$ 是弱保形的. □

由此已经可得

$$A(X_\varepsilon \circ g_0) = D(X_\varepsilon \circ g_0).$$

我们只知道对任意微分同胚 $g, A(X_\varepsilon) = A(X_\varepsilon \circ g)$, 但 $g_0 \in \bar{S}$ 未必是微分同胚. 以下要把面积的微分同胚不变性推广到 $\bar{S}$ 上去. 如其成立, 定理 17.2 的证明也就完成了. 为此, 先研究一下 $\bar{S}$.

引理 17.4 $\bar{S} \subset C(\bar{D}, \bar{D}) \cap C(\partial D, \partial D)$. $\forall g \in \bar{S}$, $g : \bar{D} \to \bar{D}$ 是满射, 并且在 ∂D 上是弱单调的, 还满足三点条件.

证明

1. 对于边界点, 利用 Courant-Lebesgue 引理, 或对于 $\Gamma = \partial D$, 直接利用引理 17.2, 即得 $\forall \varepsilon > 0, \exists \rho > 0$, 使得

$$\sup_{w, w' \in C_\rho} |g(w) - g(w')| < \varepsilon,$$

其中 $C_\rho = \partial B_\rho(w_0) \cap D$.

2. 至于内点, 因为 $g \in S$ 是微分同胚, 映 $B_\rho(w_0) \cap D$ 于由 $g(C_\rho)$ 包围的小 (拓扑) 圆域内,

$$\begin{aligned} \sup_{|w-w'|<\delta} |g(w) - g(w')| &\leqslant \sup_{w_0 \in \bar{D}} \{ \sup_{w, w' \in B_\rho(w_0) \cap D} |g(w) - g(w')| \} \\ &\leqslant \sup_{w_0 \in \bar{D}} \{ \sup_{w, w' \in C_\rho} |g(w) - g(w')| \} < \varepsilon. \end{aligned}$$

这就说明 S 的 H^1 有界子集是等度连续的.

3. $\forall g \in \bar{S}$, $\exists \{g_j\}_1^\infty \subset S$, $g_j \rightharpoonup g$ $(H^1(D))$. 前面已证 g_j 在 $\bar{D}$ 上是等度连续的. 按 Arzela-Ascoli 定理, $g_j \rightrightarrows g$ 一致收敛于 $\bar{D}$, 从而 $g \in C(\bar{D}, \bar{D})$ 是满射, 并且 $g \in C^*(\partial D)$. □

在对 X 增加可微性的情况下, 我们可以把这种不变性推广.

引理 17.5 设 $\Gamma \in C^2$ 是一条 Jordan 曲线, 而且 $\forall g \in \bar{S}, \forall X \in C^*(\Gamma) \cap C^2(\bar{D}, \mathbb{R}^n)$, 我们有 $X \circ g \in C^*(\Gamma)$, 以及 $A(X \circ g) = A(X)$.

证明 由 $X \in C^*(\Gamma)$ 以及引理 17.4, $g \in C^*(\partial D)$, 即得 $X \circ g \in C^*(\Gamma)$.

我们早已知道面积泛函在微分同胚下是不变的, 现在要做的是通过弱极限将其过渡到 $\bar{S}$ 上去. 我们分两步走.

1. 先固定 $w = (h(z)) \in \bar{S}$, $\forall g \in S$, 考察积分

$$\int_D |X_u \times X_v|_{w=(h(z))} \det(\nabla g(z)) dz.$$

注意到

$$\det(\nabla g) = \partial_x(g_1 \partial_y g_2) - \partial_y(g_1 \partial_x g_2),$$

利用分部积分, 可见

$$
\begin{aligned}
&2\int_D |X_u\times X_v|_{w=(h(z))}\det(\nabla g)dxdy\\
&=-\int_D [G_h(g_1\partial_y g_2-g_2\partial_y g_1)-H_h(g_1\partial_x g_2-g_2\partial_x g_1)]dxdy\\
&\quad+\int_{\partial D}|X_u\times X_v|_{w=(h(e^{i\theta}))}\psi'(\theta)d\theta,
\end{aligned}
$$

其中

$$\psi'(\theta)=g_1\partial_\theta g_2-g_2\partial_\theta g_1,$$

$$G_h=|X_u\times X_v|_u\partial_x h_1+|X_u\times X_v|_v\partial_x h_2,$$

$$H_h=|X_u\times X_v|_u\partial_y h_1+|X_u\times X_v|_v\partial_y h_2.$$

因为 $G_h,H_h\in L^2(D)$, 所以等式右端面积分可以连续扩张到 $g\in\bar S$.

在等式右边的线积分中, 我们注意到

$$g_1^2+g_2^2=1,$$

因此

$$\psi(\theta)=\arctan\frac{g_2(e^{i\theta})}{g_1(e^{i\theta})},$$

当 $g\in\bar S$ 时, $g\in C(\bar D,\bar D)$, 而且 $\partial D\to\partial D$ 是单调的, 即 ψ 是单调的, 所以 $\psi'(\theta)$ 几乎处处存在. 此外,

$$\int_{\partial D}|\psi'(\theta)|d\theta=2\pi,$$

可见 $\psi\in W^{1,1}(\partial D)$. 这表明等式右边的线积分也可以连续扩张到 $g\in\bar S$. 即若 $\{g^j\}\subset S,\ g\in\bar S$, 使得 $g^j\rightharpoonup g$, 则

$$
\begin{aligned}
&2\int_D |X_u\times X_v|_{w=(g(z))}\det(\nabla g)dxdy\\
&=\lim_{j\to\infty}\Big[-\int_D [G_g(g_1^j\partial_y g_2^j-g_2^j\partial_y g_1^j)-H_g(g_1^j\partial_x g_2^j-g_2^j\partial_x g_1^j)]dxdy\\
&\quad+\int_{\partial D}|X_u\times X_v|_{w=(g(e^{i\theta}))}\psi_j'(\theta)d\theta\Big]\\
&=\lim_{j\to\infty}2\int_D |X_u\times X_v|_{w=(g(z))}\det(\nabla g^j)dxdy,
\end{aligned}
$$

其中 $\psi'(\theta)=g_1^j\partial_\theta g_2^j-g_2^j\partial_\theta g_1^j$.

2. 我们再证明: 若 $\{g_j\} \subset S$, $g_j \rightharpoonup g\,(H^1 \cap C(\bar{D}, \bar{D}))$, 则

$$\int_D \Big| |X_u \times X_v|_{w=(g_j(z))} - |X_u \times X_v|_{w=(g(z))} \Big| \det(\nabla g_j) dxdy \to 0.$$

一方面, 这是因为引理 17.4 蕴含了 $g_j \rightrightarrows g$ 一致收敛于 $\bar{D}$, 所以

$$\Big| |X_u \times X_v|_{w=(g_j(z))} - |X_u \times X_v|_{w=(g(z))} \Big| \rightrightarrows 0.$$

另一方面,

$$\int_D |\det(\nabla g_j)| dxdy \leqslant \int_D |\nabla g_j|^2 dxdy \leqslant M,$$

即得结论.

3. 现在直接由

$$A(X \circ g) = A(X), \quad \forall g \in S$$

推出

$$A(X \circ g) = A(X), \quad \forall g \in \bar{S}.$$

□

定理 17.2 的证明 对于有可微性假设的 Jordan 曲线 Γ, 用它作为边值求解调和方程, 就得到 $C(\Gamma) \neq \varnothing$.

从前面的分析, 我们只要证明 $\inf D(X) \leqslant \inf A(X)$ 就够了. 联合引理 17.3 与 17.5, 对于 $X \in C^*(\Gamma) \cap C^2(\bar{D}, \mathbb{R}^n), \exists g \in \bar{S}$, 使得

$$A(X_\varepsilon) = A(X_\varepsilon \circ g) = D(X_\varepsilon \circ g).$$

我们知道

$$D(X \circ g) \leqslant D(X_\varepsilon \circ g),$$

以及

$$\lim_{\varepsilon \to 0} A(X_\varepsilon) = A(X),$$

于是

$$\inf_{X \in C(\Gamma)} D(X) \leqslant \inf_{X \in C(\Gamma) \cap C^2(\bar{D}, \mathbb{R}^n)} A(X).$$

又因为假定了 $\Gamma \in C^2$, 所以 $C^2(\bar{D}, \mathbb{R}^n) \cap C(\Gamma)$ 在 $C(\Gamma)$ 中是稠密的, 定理证完.

□

第十八讲　变分问题的数值方法

§18.1　Ritz 方法

变分问题的直接解法由两部分组成:

(1) 构造一串极小化序列,

(2) 对极小化序列取极限获得问题的解.

L. Rayleigh 和 W. Ritz 提出了一种数值求解极小化序列的方法, 称为 Rayleigh-Ritz 方法. 设 M_0 是区域 Ω 上函数空间 X 的一个线性子空间, I 是 Ω 上的积分泛函 $\int_\Omega L(x, u(x), \nabla u(x))\, dx$, 其定义域 $M = \varphi_0 + M_0$, φ_0 是 Ω 上的一个函数 (由非齐次边值确定).

Rayleigh-Ritz 方法的思想是选取 M_0 的一组完备基 $\{e_1, e_2, \cdots\}$, 令

$$M_n = \varphi_0 + \operatorname{span}\{e_1, e_2, \cdots, e_n\} \subset M.$$

我们在 M_n 上求泛函 I 的极小点 $\varphi_n \in M_n$. 因为 M_n 上任意函数可以写成

$$u = \varphi_0 + \sum_{i=1}^{n} \xi_i e_i, \quad \xi = (\xi_1, \cdots, \xi_n) \in \mathbb{R}^n \tag{18.1}$$

的形式, 所以 $I|_{M_n}$ 实际上是 n 个实变量 ξ 的函数:

$$I(u) = I\left(\varphi_0 + \sum_{i=1}^{n} \xi_i e_i\right). \tag{18.2}$$

我们可以直接对 $I|_{M_n}$ 求极小值 (优化技术), 或者通过求解 n 个齐次方程组:

$$\frac{\partial}{\partial \xi_j} I\left(\varphi_0 + \sum_{i=1}^{n} \xi_i e_i\right) = 0, \qquad j = 1, 2, \cdots, n, \tag{18.3}$$

来得到 φ_n.

与这种想法密切关联的是 Ritz-Galerkin 方法, 这个方法并不针对泛函 I, 而是针对它的 E-L 方程:

$$\delta I(u, \varphi) = \int_{\Omega} (L_u(\tau)\varphi(x) + L_p(\tau)\nabla\varphi(x))\, dx = 0, \quad \forall \varphi \in M_0,$$

其中 $\tau = (x, u(x), \nabla u(x))$. 在 M_n 上还是把 u 表示成 (18.1) 的形式, 逐次选取 $\varphi = e_j, j = 1, \cdots, n$, 我们就能得到

$$\int_{\Omega} (L_u(\tau_n(x))e_j(x) + L_p(\tau_n(x))\nabla e_j(x))\, dx = 0, \quad j = 1, \cdots, n, \tag{18.4}$$

其中

$$\tau_n(x) = \left(x, \varphi_0(x) + \sum_{i=1}^{n} \xi_i e_i(x), \nabla\varphi_0(x) + \sum_{i=1}^{n} \xi_i \nabla e_i(x)\right).$$

这时, (18.4) 还是一个 n 个变量 $\xi = (\xi_1, \cdots, \xi_n)$ 的 n 个方程的方程组.

例 18.1 设 $f \in L^2(\Omega)$, 定义泛函

$$I(u) = \int_{\Omega} \left[\frac{1}{2}|\nabla u|^2 - f \cdot u\right] dx, \quad u \in H_0^1(\Omega). \tag{18.5}$$

取 V_n 为 $H_0^1(\Omega)$ 的 n 维线性子空间. 按 Rayleigh-Ritz 方法, 化归为求函数

$$\begin{aligned} J_n(\xi_1, \cdots, \xi_n) &= \int_{\Omega} \left[\frac{1}{2}\sum_{i=1}^{n} |\xi_i \nabla e_i(x)|^2 - f(x)\sum_{i=1}^{n} \xi_i e_i(x)\right] dx \\ &= \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}\xi_i\xi_j - \sum_{i=1}^{n} b_i\xi_i \end{aligned}$$

的极小值, 其中 $a_{ij} = \displaystyle\int_{\Omega} \nabla e_i(x)\nabla e_j(x)\, dx, i, j = 1, \cdots, n$, $b_j = \displaystyle\int_{\Omega} f(x)e_j(x)\, dx$, $j = 1, \cdots, n$.

或者等价地求解下列线性方程组:

$$\sum_{j=1}^{n} a_{ij}\xi_j = b_i, \quad i = 1, 2, \cdots, n. \tag{18.6}$$

再来看相应的 E-L 方程

$$\int_\Omega [\nabla u\cdot\nabla\varphi - f\cdot\varphi]\,dx = 0,\quad \forall\varphi\in H_0^1(\Omega).$$

按 Ritz-Galerkin 方法, 我们要求解方程组

$$\sum_{i=1}^n \xi_j\int_\Omega \nabla e_i(x)\nabla e_j(x)\,dx - \int_\Omega f(x)e_j(x)\,dx = 0,\quad j=1,\cdots,n,$$

即

$$\sum_{i=1}^n a_{ij}\xi_i - b_j = 0,\quad j=1,\cdots,n. \tag{18.7}$$

注意到 $a_{ij}=a_{ji},\forall i,j$, 方程组 (18.6) 与 (18.7) 是一样的. □

注 18.1　Ritz-Galerkin 方法不仅能应用于泛函的 E-L 方程, 而且还能应用到一切通过积分形式定义弱解的微分方程. 例如, 设 a_{ij},b_j,c 都是连续函数, 人们把满足方程

$$\begin{cases}\displaystyle\sum_{i,j=1}^n \frac{\partial}{\partial x_j}\left(a_{ij}(x)\frac{\partial}{\partial x_j}u(x)\right)+\sum_{j=1}^n b_j(x)\frac{\partial}{\partial x_j}u(x)+c(x)u(x)=f(x), & x\in\Omega,\\ u|_{\partial\Omega}=0 & (18.5')\end{cases}$$

的弱解定义为 $\forall\varphi\in H_0^1(\Omega)$, 满足:

$$\int_\Omega\left[-\sum_{i,j=1}^n a_{ij}(x)\frac{\partial u(x)}{\partial x_j}\frac{\partial\varphi(x)}{\partial x_i}+\left(\sum_{j=1}^n b_j(x)\frac{\partial u(x)}{\partial x_j}+c(x)u(x)-f(x)\right)\varphi(x)\right]dx=0$$

的 $u\in H_0^1(\Omega)$.

我们照样可以用 $u=\sum_{k=1}^n\xi_k e_k(x)$, $\varphi(x)=e_l(x), l=1,2,\cdots,n$, 代入得到 n 个变元的 n 线性方程组. □

§18.2　有限元

那么怎样构造有限维逼近子空间呢? 我们提出一个原则, 按照这个原则构造出一列有限维子空间, 使得它们可以任意逼近给定的空间. 以变分问题中经常遇到的空间 $H_0^1(\Omega)$ (Dirichlet 零边值问题) 和空间 $H^1(\Omega)$ (Neumann 零边值) 为例, 我们按下列方式构造 Ω 上的函数空间作为上述空间的有限维子空间.

为简单计, 假定区域 Ω 是一个有界的多面体, 换句话说, $\partial\Omega$ 是由逐片超平面构成的. 我们选取逐片线性函数组成的空间来逼近 $H_0^1(\Omega)$ (或 $H^1(\Omega)$).

我们先引进几个术语. 称 $K \subset \mathbb{R}^n$ 是一个 n-单形, 记作 $K=\{P_0,P_1,\cdots,P_n\}$, 如果它是这 $n+1$ 个点 $p_0,\cdots,p_n$ 的闭凸包, 其中 $P_i \in \mathbb{R}^n, i=0,1,\cdots,n$, 使得

$$\{P_i - P_0 \mid 1 \leqslant i \leqslant n\}$$

是线性无关的向量.

因此 0-单形是点, 1-单形是线段, 2-单形是三角形, 如此等等.

称 $J=\{K_1,\cdots,K_m\}$ 是 Ω 的一个三角剖分, 如果

(1) $K_i \cap K_j = \varnothing$ 或一个 p-单形 $0 \leqslant p \leqslant n-1$,

(2) $\Omega = \bigcup_{K_i \in J} K_i$.

记 $h_J = \max\{\mathrm{diam}(K_i) \mid 1 \leqslant i \leqslant m\}$, 称之为 J 的尺寸.

对空间 $H^1(\Omega)$, $\{K_i\}_1^m$ 的顶点都称为结点(node), 记作 $N=\{N_j \mid 1 \leqslant j \leqslant M\}$.

对空间 $H_0^1(\Omega)$, 则舍去在边界 $\partial\Omega$ 上的顶点, 只取在 Ω 内的顶点, 记作 $\{N_j \mid 1 \leqslant j \leqslant M_0\}$.

我们将所有结点对应着有限维逼近子空间 V_h 的对偶基.

选取 V_h 中的元素为逐片仿射函数, 其基函数为 $\{e_i\}_1^M$, 它们由结点集决定,

$$e_i(N_j) = \delta_{ij}, \qquad i,j=1,2,\cdots,M.$$

事实上, 对于任意单形 $K \in J$, 基函数 e_i 在 K 的值完全由它在 K 的各个顶点上的值决定 (当顶点在 $\partial\Omega$ 上时, 对齐次 Dirichlet 边值取零值, 对非齐次边值则取给定的值).

特别地, 取 $K=(P_0,\cdots,P_n)$, $P_i=(p_i^1,\cdots,p_i^n), i=0,1,\cdots,n$, 并引入 $n+1$ 个函数

$$\begin{cases} v_j(x) = x^j, \quad j=1,2,\cdots,n, \\ v_0(x) = 1, \end{cases}$$

其中 $x=(x^1,\cdots,x^n)$, 便有

$$\begin{cases} x^j = \displaystyle\sum_{i=0}^n p_i^j e_i(x), \\ 1 = \displaystyle\sum_{i=0}^n e_i(x). \end{cases} \tag{18.8}$$

由此可见, $\forall i$, e_i 的支集是 $\bigcup_{K\in J}\{K \mid N_i \in K\}$, 即那些以 N_i 为其顶点的单形之并集.

V_h 中任意函数 $v \in V_h$ 有表示

$$v(x) = \sum_{j=1}^{M} v(N_j) e_j(x).$$

我们称 (Ω, V_h, N) 为一个有限元 (图 18.1).

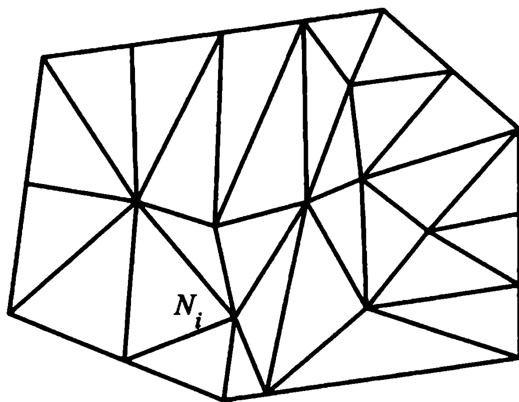


图 18.1

任意给定一个 $C(\bar{\Omega})$ 函数 u, 怎样构造 V_h 中的元素与之逼近? 最自然的也是最简单的办法是插值.

记 Πu 为 u 在 V_h 中的对应元素. $\forall K \in J$, 记 $\Pi_K u = \Pi u|_K$. 因为 V_h 中的函数都是逐片仿射的, 所以 $\Pi_K u$ 由 u 在 K 的顶点上的值完全确定. 令

$$\Pi_K u(x) = \sum_{j=0}^{n} u(P_j) e_j(x),$$

其中 $K = \{P_0, \cdots, P_n\}$.

由此可见, $\Pi : C(\bar{\Omega}) \to H_0^1(\Omega)$ 是一个有界线性算子, 并且

$$\|\Pi u\|_{H^1} \leqslant C \|u\|_{C(\bar{\Omega})},$$

其中 $C \doteq n \sup_{1 \leqslant j \leqslant M} \|e_j\|_{H^1}$.

引理 18.1 设 $u \in C^2(\bar{\Omega})$, 则

$$\|\nabla u - \nabla(\Pi u)\|_{1,\infty} \leqslant \frac{n^2(n+1)}{2} \frac{h_J^2}{\rho_J} \|\nabla^2 u\|_\infty, \tag{18.9}$$

$$\|u - \Pi u\|_\infty \leqslant \frac{n^2(n+1)}{2} h_J^2 \|\nabla^2 u\|_\infty, \tag{18.10}$$

其中 $\|\cdot\|_{1,\infty}$ 是 $W^{1,\infty}$ 模, $\|\cdot\|_\infty$ 是 L^∞ 模,

$$h_J = \sup_{K \in J} h_K, \quad h_K = \operatorname{diam}(K),$$

以及

$$\rho_J = \inf_{K \in J} \rho_K, \quad \rho_K = \sup\{2R \mid B_R(x) \subset K,\, x \in K\}.$$

证明　我们只需要在每一个单形 $K = (P_0, P_1, \cdots, P_n)$ 上考察 $\nabla u(x)$ 与 $\nabla(\Pi_K u)$ 之差就够了. 对 $\nabla(\Pi_K u)(x)$ 作 Taylor 展开:

$$\begin{aligned}\nabla(\Pi_K u)(x) &= \sum_{j=0}^{n} u(P_j)\nabla e_j(x)\\ &= \sum_{j=0}^{n}\left[u(x) + \nabla u(x)(P_j - x) + \frac{1}{2}\nabla^2 u(\xi_x)(P_j - x)^2\right]\nabla e_j(x),\end{aligned} \tag{18.11}$$

其中 $\xi_x \in K$. 对 (18.8) 作微商, 我们得到

$$\begin{cases} \delta_k^j = \displaystyle\sum_{i=0}^{n} p_i^j \frac{\partial e_i}{\partial x_k}, \\ 0 = \displaystyle\sum_{i=0}^{n} \frac{\partial e_i}{\partial x_k}. \end{cases} \qquad j, k = 1, 2, \cdots, n,$$

代入 (18.11) 得到

$$\nabla(\Pi_K u)(x) = \nabla u(x) + \sum_{j=0}^{n} R_j(x)\nabla e_j(x),$$

其中

$$R_j(x) = \frac{1}{2}\nabla^2 u(\xi_x)(P_j - x)^2.$$

由于

$$\|P_j - x\|_{\mathbb{R}^n} \leqslant \mathrm{diam}K \leqslant h_J, \qquad \left|\frac{\partial e_i}{\partial x_k}\right| \leqslant \frac{1}{\rho_K} \leqslant \frac{1}{\rho_J},$$

就得到 (18.9),

$$\|\nabla u - \nabla(\Pi u)\|_{1,\infty} \leqslant \frac{n^2(n+1)}{2}\frac{h_J^2}{\rho_J}\|\nabla^2 u\|_\infty.$$

同理可证 (18.10).　□

定理 18.1　设 $u \in C(\bar{\Omega}) \cap H^2(\Omega)$, 又设对单形剖分 J, 存在 $\beta > 0$, 使得

$$\frac{\rho_K}{h_K} \geqslant \beta, \quad \forall K \in J, \tag{18.12}$$

则

$$\|u - \Pi u\|_{L^2} \leqslant C h_J^2 |u|_{H^2},$$

以及

$$|u-\Pi u|_{H^1} \leqslant C_\beta h_J |u|_{H^2},$$

其中半模

$$|v|_{H^r} = \left(\sum_{|\alpha|=r} \int_\Omega |D^\alpha v|^2\, dx\right)^{1/2}, \quad r=1,2.$$

证明 按引理 18.1 对于任意一个单形 $K \in J$, 我们有

$$\|u-\Pi u\|_{L^2(K)} \leqslant C h_K^2 |u|_{H^2(K)},$$

以及

$$|u-\Pi u|_{H^1(K)} \leqslant C\frac{h_K^2}{\rho_K}|u|_{H^2(K)}.$$

从而

$$\begin{aligned}
\|u-\Pi u\|_{L^2(\Omega)}^2 &= \sum_{K\in J} \|u-\Pi_K u\|_{L^2(K)}^2 \\
&\leqslant \sum_{K\in J} C^2 h_K^4 |u|_{H^2(K)}^2 \\
&\leqslant C^2 h_J^4 |u|_{H^2}^2,
\end{aligned}$$

以及

$$\begin{aligned}
|u-\Pi u|_{H^1(\Omega)}^2 &= \sum_{K\in J} |u-\Pi_K u|_{H^1(K)}^2 \\
&\leqslant \sum_{K\in J} C^2 \frac{h_K^4}{\rho_K^2} |u|_{H^2(K)}^2 \\
&\leqslant C^2 h_J^2 \frac{1}{\beta^2} |u|_{H^2(K)}^2.
\end{aligned}$$ □

注 18.2 条件 (18.12) 的几何意义是: 在单形剖分中, 每个单形 K 的角度不能任意小. □

注 18.3 无论在理论研究, 还是在实际计算上, 采用插值方法构造逼近函数好处较多. 原因在于用插值法对于解函数只要求是连续的, 即 $u \in C(\bar\Omega)$, 而得到的却是 H^1 意义下的逼近. 对于正则变分问题, 或是对于解有正则性的微分方程, 解总是连续的. 特别地, 对于二阶线性椭圆型方程, $u \in H^2(\Omega)$. 当 $n \leqslant 3$ 时, 由嵌入定理得 $H^2(\Omega) \hookrightarrow C(\bar\Omega)$. □

§18.3 Cea 定理

我们从一个抽象定理出发. 设 H 是一个 Hilbert 空间, 有内积 $(\cdot,\cdot)$ (对应的模是 $\|\cdot\|$). 又设 V 是 H 的一个闭线性子空间. 给定 $F \in V^*$.

假设

(1) $a(u,v)$ 是 H 上的一个有界、双线性形式, 即

$$|a(u,v)| \leqslant C\|u\|_H\|v\|_H,$$

$$a(\alpha_1 u_1 + \alpha_2 u_2, v) = \alpha_1 a(u_1,v) + \alpha_2 a(u_2,v),$$

$$a(u, \beta_1 v_1 + \beta_2 v_2) = \beta_1 a(u,v_1) + \beta_2 a(u,v_2).$$

(2) a 在 V 上是强制的, 即 $\exists\,\alpha > 0$,

$$a(v,v) \geqslant \alpha\|v\|^2, \quad \forall v \in V.$$

(3) V_h 是 V 的一个有限维闭子空间.

按 Lax-Milgram 定理, 在假设 (1)、(2) 之下, 存在唯一的 $u \in V$, 使得

$$a(u,v) = F(v), \quad \forall v \in V.$$

又按 Ritz-Galerkin 方法得到逼近解, 存在唯一的 $u_h \in V_h$, 适合

$$a(u_h,v) = F(v), \quad \forall v \in V_h.$$

问 u_h 在什么意义下逼近 u?

定理 18.2 (Cea) *在假设 (1)–(3) 下, 若 $u \in V$ 是*

$$a(u,v) = F(v), \quad \forall v \in V$$

的唯一解, 又若 $u_h \in V_h$ 是近似方程

$$a(u_h,v) = F(v), \quad \forall v \in V_h$$

的解, 则

$$\|u - u_h\| \leqslant \frac{C}{\alpha}\min_{v\in V_h}\|u - v\|.$$

证明 由于

$$a(u,v)=F(v),\qquad \forall v\in V,$$
$$a(u_h,v)=F(v),\qquad \forall v\in V_h,$$

所以

$$a(u-u_h,v)=0,\quad \forall v\in V_h.$$

特别地,

$$\begin{aligned}\alpha\|u-u_h\|^2&\leqslant a(u-u_h,u-u_h)\\&=a(u-u_h,u-v)+a(u-u_h,v-v_h)\\&\leqslant C\|u-u_h\|\,\|u-v\|,\end{aligned}$$

其中 $v\in V_h$.

这就是

$$\|u-u_h\|\leqslant\frac{C}{\alpha}\|u-v\|,\quad \forall v\in V_h,$$

即

$$\|u-u_h\|\leqslant\frac{C}{\alpha}\min_{v\in V_h}\|u-v\|.$$ □

现在回到注 18.1 中具体的线性二阶椭圆型方程 (18.5′). 假设 $\alpha,\beta,\gamma>0$, 满足 $\beta^2<\alpha\gamma$, 以及

(1) $\sum_{i,j=1}^n a_{ij}(x)\xi_i\xi_j\geqslant\alpha|\xi|^2$, $\forall\xi=(\xi_1,\cdots,\xi_n)\in\mathbb{R}^n$, $|a_{ij}(x)|\leqslant C$, $1\leqslant i,j\leqslant n$, $\forall x\in\Omega$,

(2) $c(x)\geqslant\gamma$, $\forall x\in\Omega$,

(3) $\max\limits_{1\leqslant i\leqslant n}\|b_i\|_\infty<2\beta$.

令 $\delta=\alpha-\frac{\beta^2}{\gamma}$, 则 $\delta>0$. 又令

$$a(u,v)=\int_\Omega\left(\sum_{i,j=1}^n a_{ij}(x)\partial_j u\,\partial_i v+\sum_{i=1}^n b_i(x)\partial_i u\cdot v+c(x)u\cdot v\right)dx,\qquad u,v\in H_0^1(\Omega).$$

如此定义的 $a(u,v)$ 是 $H_0^1(\Omega)$ 上的一个有界双线性形式. 从不等式

$$\left|\int_\Omega\sum_{i=1}^n b_i(x)\partial_i u(x)\cdot u(x)\,dx\right|\leqslant\max_{1\leqslant i\leqslant n}\|b_i\|_\infty\|\nabla u\|_2\|u\|_2$$

可见双线性形式 $a(u,v)$ 在 $H_0^1(\Omega)$ 上是强制的:

$$a(v,v) \geqslant \alpha\|\nabla v\|^2 - 2\beta\|\nabla v\|\,\|v\|_2 + \gamma\|v\|_2^2 \geqslant \delta\|\nabla v\|^2.$$

如今 $\forall f \in L^2(\Omega)$, 应用 Lax-Milgram 定理, 我们得到唯一的解 $u \in H_0^1(\Omega)$, 满足

$$a(u,v) = \int_\Omega f \cdot v\,dx, \quad \forall\, v \in H_0^1(\Omega).$$

又利用椭圆型方程的正则性, 可以知道 $u \in H^2(\Omega)$, 并且存在常数 C 满足:

$$|u|_{H^2(\Omega)} \leqslant C\|f\|_{L^2(\Omega)}.$$

如果我们用 Ritz-Galerkin 方法近似求解时, 采用前面所说的有限元 (Ω, V_h, N), 得到 $u_h \in V_h$, 那么再根据 Cea 定理以及定理 18.1,

$$\begin{aligned}\|u-u_h\|_{H_0^1(\Omega)} &= \frac{C}{\delta}\min_{v\in V_h}\|u-v\|_{H_0^1(\Omega)}\\ &\leqslant \frac{C}{\delta}\,h_J|u|_{H^2(\Omega)}\\ &\leqslant \frac{C'}{\delta^2}\,h_J\|f\|_{L^2(\Omega)}.\end{aligned}$$

这表明采用有限元方法, 不仅插值函数在 H_0^1 模意义下逼近原来的函数, 而且作为近似解 u_h 也可以在 H_0^1 模意义下逼近真解 u. 其逼近的程度与网格的尺寸 h_J 成正比.

§18.4 最优化方法 —— 共轭梯度法

最优化技术是一种数值求解函数极值的技术. 在 $\mathbb{R}^n$ 上给定一个实函数 $f:\mathbb{R}^n \to \mathbb{R}^1$, 要设计出一种迭代算法确定一个极小化序列 $x^0, x^1, \cdots, x^i, \cdots$, 把函数值引向极小值.

从一点 x^i 出发, 为了每迭代一步函数值都能下降, $f(x^{i+1}) < f(x^i)$(如果可能的话), 我们应当怎样选择 x^{i+1}?

注意到, 从一点 x^i 到另一点 x^{i+1} 的移动有两个要素——方向和步长,

$$x^{i+1} - x^i = \lambda^i\sigma^i, \quad \lambda^i > 0,\ \|\sigma^i\| = 1, \tag{18.13}$$

其中 σ^i 表示下降的方向, 而 λ^i 表示步长 (图 18.2).

在点 x^i 附近将 f 线性化:

$$f(x) = f(x^i) + (\nabla f(x^i), x - x^i) + o(\|x - x^i\|).$$

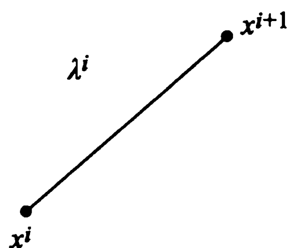


图 18.2 $\lambda^i = \|x^{i+1} - x^i\|, \sigma^i = \frac{x^{i+1}-x^i}{\|x^{i+1}-x^i\|}$

局部地看, 选择负梯度方向

$$\sigma^i = -\nabla f(x^i)/\|\nabla f(x^i)\| \tag{18.14}$$

是最速下降方向. 步长 λ^i 可以通过求沿这个方向的一元函数

$$\phi(t) = f(x^i + t\sigma^i) \tag{18.15}$$

的极小值得到. 这样形成的算法称为梯度法, 其步骤如下.

1. 给定初始点 x^0.
2. 从 x^i 出发, 按 (18.14) 式算出 σ^i, 再按
$$\begin{cases} \phi'(\lambda^i) = 0, \\ \phi(\lambda^i) < \phi(\lambda), \quad \forall \lambda \in [0, \lambda^i), \end{cases}$$
计算出 $\lambda^i > 0$, 其中 $\phi(t)$ 由 (18.15) 式定义.
3. 按 (18.13) 式, 由 x^i, σ^i 与 λ^i 确定 x^{i+1}.
若 $\sigma^{i+1} = 0$, 则停止, x^{i+1} 即为所求;
否则继续下去.

不难证明, 如果 f 是强制的, 而且只有一个极小点, 那么从任意初始点 x^0 出发, 按照梯度法或者经过有限步达到极小点, 或者得到一串极小化子序列收敛到这个极小点.

但梯度法并不十分理想, 原因是有时收敛速度非常慢. 从下面一个例子就可以看出其原因所在. 设

$$f(x_1, x_2) = x_1^2 + \frac{1}{\varepsilon} x_2^2, \quad \varepsilon > 0.$$

这是一个等高线都是椭圆的二次函数, 其极小点在 (0,0). 给定 $x^i = (x_1^i, x_2^i)$, 则推出

$$x^{i+1} = \left((1 + \lambda^i)x_1^i, \left(1 + \frac{\lambda^i}{\varepsilon}\right) x_2^i\right),$$

其中

$$\lambda^i = \frac{-\left((x_1^i)^2 + \frac{1}{\varepsilon^2}(x_2^i)^2\right)}{(x_1^i)^2 + \frac{1}{\varepsilon^3}(x_2^i)^2}.$$

当 $\varepsilon > 0$ 很小时,

$$x^{i+1} \approx \left((1-\varepsilon)x_1^i, -\varepsilon^2 \frac{(x_1^i)^2}{x_2^i}\right).$$

可见下降路径来回曲折缓慢地接近极小点 (0,0) (图 18.3).

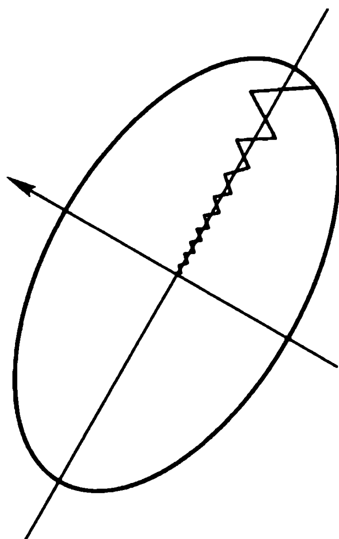


图 18.3

共轭梯度法是为了克服梯度法上述缺点提出来的一种算法. 其基本思想是: 在极小点的一个小邻域内, 函数 f 可以用二阶 Taylor 展开近似表示:

$$f(x) = f(x_0) + (\nabla f(x_0), x - x_0) + \frac{1}{2}\left(\nabla^2 f(x_0)(x - x_0), x - x_0\right) + o(\|x - x_0\|^2).$$

对于一个有唯一极小点的二次函数

$$c + (b, x) + \frac{1}{2}(Ax, x), \quad c \in \mathbb{R}^1,\ b \in \mathbb{R}^n,\ A \in \mathbb{R},$$

矩阵 A 是正定的. 当 $n = 2$ 时, 它就是以一族椭圆为等高线的函数. 我们将充分利用这个二次函数的特性来避免来回曲折的下降路径.

从解析几何知道, 一个给定的平面向量 $w = (w_1, w_2)$ 决定了这族椭圆的一条直径. 而和它共轭的直径正是与这方向平行的诸弦的中点的轨迹. 我们把这个共轭直径的方向 u 称为 w 的共轭方向 (图 18.4).

这个几何事实启发我们: 对于平面上的正定二次函数, 不论从哪一点 x^0 出发, 任取一个方向 w, 沿 w 取极小点, 这点应在椭圆的平行于 w 的弦的中点上. 下一步

只要再沿与 w 共轭的方向 u 求极小点, 两步就达到了这族椭圆的中心, 也就是原来函数的极小点.

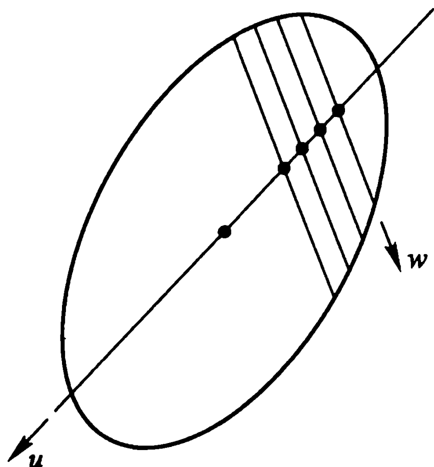


图 18.4

为了把上述几何事实推广到高维, 我们先把"共轭"的概念解析地表达出来. 为简单计, 取 $b = 0$.

给定向量 w, $\forall v \in \mathbb{R}^2$, 过点 v 平行于 w 的直线的参数方程是 $x = v + tw$. 它与椭圆

$$c + \frac{1}{2}(Ax, x) = 0$$

的交点是

$$x_i = v + t_i w, \quad i = 1, 2,$$

其中 t_1, t_2 是方程

$$2c + t^2(Aw, w) + 2t(Av, w) + (Av, v) = 0$$

的两个实根.

如果 v 在共轭直径上, 那么 v 就是 $\frac{x_1+x_2}{2}$. 从而

$$t_1 + t_2 = 0,$$

亦即

$$(Av, w) = 0.$$

换句话说, 如果我们按正定矩阵 A 重新规定 $\mathbb{R}^2$ 上的尺度 (内积):

$$[v, w]_A = (Av, w),$$

那么

$$v \text{ 与 } w \text{ 共轭} \iff v \text{ 按内积 } [\cdot, \cdot]_A \text{ 与 } w \text{ 正交}.$$

这件事容易推广到$\mathbb{R}^n$. 设 A 是一个$n\times n$ 正定矩阵, $\forall v,w\in\mathbb{R}^n$, 定义

$$[v,w]_A=(Av,w),$$

容易验证, $[\cdot,\cdot]_A$ 是一个内积.

我们称 v 与 w 按 A 共轭, 如果

$$[v,w]_A=(Av,w)=0.$$

因此在$\mathbb{R}^n$ 上, 对于给定的正定矩阵 A, 存在着n 个相互共轭的向量组 $w_1,\cdots,w_n$,

$$(Aw_i,w_j)=0,\quad 1\leqslant i<j\leqslant n.$$

给定一个二次函数$f(x)=c+(b,x)+\frac{1}{2}(Ax,x)$, 其中 A 是正定的. 作为一个线性代数的习题, 读者不难验证: 从任意点x^0 出发, 寻找一组相互共轭的单位向量$w_1,\cdots,w_n$, 逐次沿这些单位向量求一维极小点, 求n 步就能达到f 的极小点.

根据这个思想构造出来的共轭梯度算法如下.

1. 给定初始点$x^0\in\mathbb{R}^n$ 与初始方向$\sigma^0=-\nabla f(x^0)$.

2. 设已有x^i 与σ^i, 令

$$\begin{aligned}\lambda^i&=-(r^i,\sigma^i)/(A\sigma^i,\sigma^i),\\ \sigma^{i+1}&=-r^{i+1}+\beta^i\sigma^i,\\ \beta^i&=(Ar^{i+1},\sigma^i)/(A\sigma^i,\sigma^i),\end{aligned}$$

   其中

$$r^i=\nabla f(x^i).$$

   取

$$x^{i+1}=x^i+\lambda^i\sigma^i.$$

3. 若$r^{i+1}=0$, 则终止, 已得到极小点x^{i+1}; 否则继续下去. 对于二次函数, 上述步骤至多进行n 步必将终止.

事实上不难用数学归纳法验证: 当$m\leqslant n$ 时,

(1) 若$\phi(t)=f(x^i+t\sigma^i)$, 则$\phi'(\lambda^i)=0$.

(2) $\mathrm{span}\{\sigma^0,\cdots,\sigma^m\}=\mathrm{span}\{r^0,\cdots,r^m\}=\mathrm{span}\{r^0,Ar^0,\cdots,A^mr^0\}$.

(3) $[\sigma^i, \sigma^j]_A = 0, \quad i \neq j, \quad 1 \leqslant i < j \leqslant m.$

注 18.4 因为二次函数优化问题可以化归线性方程组求解, 所以共轭梯度法同样是求解线性方程组的一种有效的迭代方法.

有限元方法导出的线性方程组对应着稀疏矩阵, 采用共轭梯度法迭代求解非常有效. □

注 18.5 共轭梯度法不仅适用于正定二次函数的优化, 对于凸函数也可以用, 当然不是 n 步就能到位, 但收敛速度比梯度算法有很大提高. □

注 18.6 在数学物理, 特别是量子力学计算中, 经常要计算正定自伴算子的谱. 在一定的紧性假设下, 谱由特征值组成. 特征值有变分刻画 (见第十二讲), 从而可以应用 Ritz 方法近似求解. 有限元与共轭梯度法都是应用 Ritz 方法时的数值手段. □

第十九讲　最优控制问题

最优控制问题是一类特殊的变分问题, 在工程、经济等领域有广泛的应用. 因为实现泛函极值的函数一般不是连续的 (经常是跳跃的, switching), 所以不能用 Euler-Lagrange 方程求解, 取而代之的必要条件是 Pontryagin 极大值原理. 最优控制理论已发展成一个专门的学科, 在本讲义中我们只能作最简单的介绍.

§19.1　问题的提法

为了使读者对于最优控制问题的提法和可能的解法有一个直观的了解, 我们从一个非常简单的具体的例子出发.

例 19.1 (滚动小车问题)　把质量为 $m=1$ 的小滚动车放在 x 坐标轴上, 从静止点 $x=0$ 出发, 要想恰好走时间 T 达到点 $x=L$, 问在不考虑摩擦力与空气阻力的前提下, 应怎样加非负的外力 $u\geqslant 0$, 使得所做的功 W 最小?

用 $x(t)$ 表示小车在时刻 t 的位置, $v(t)$ 是此刻小车的速度, 用 $u(t)$ 表示在时刻 t 小车所受的外力. 它们之间关系如下:

$$\begin{cases} \dfrac{dx}{dt}=v, \\ \dfrac{dv}{dt}=u, \end{cases} \tag{19.1}$$

$$\begin{pmatrix} x(0) \\ v(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{19.2}$$
$$x(T) = L.$$

可控制的外力属于集合

$$U = \{u \in L^{\infty}([0,T]) \mid u \geqslant 0\}.$$

外力所做的功

$$W = \int_0^L u\,dx = \int_0^T uv\,dt = \int_0^T \dot{v}v\,dt = \frac{v^2(T)}{2}.$$

我们要求解约束极小化问题

$$\min\left\{W = \int_0^T uv\,dt \middle| \frac{dx}{dt} = v, \frac{dv}{dt} = u, x(0) = 0, v(0) = 0, \int_0^T v\,dt = L, u \in U\right\}.$$

分析　1° 在指定的约束条件下, $W \geqslant \dfrac{L^2}{2T^2}$.

事实上 v 是连续的, 根据积分中值定理, $\exists t_0 \in (0,T)$ 使得 $v(t_0) = L/T$. 由于 $u \geqslant 0$, v 是不减的, 所以

$$W = \frac{v^2(T)}{2} \geqslant \frac{v^2(t_0)}{2} = \frac{L^2}{2T^2}.$$

2° 我们进一步断言: $W > \frac{L^2}{2T^2}$.

事实上, 若存在控制 u 使得 $W = \frac{L^2}{2T^2}$, 那么必然存在 $t_0 \in (0,T)$ 满足

$$\begin{cases} v(t) = v(t_0) = L/T, & \forall t \in [t_0, T], \\ v(t_0) = \max\limits_{t \in [0,t_0]} v(t). \end{cases}$$

由此可得

$$L = \int_0^{t_0} v(t)\,dt + \frac{L}{T}(T - t_0),$$

即

$$\frac{L}{T}t_0 = \int_0^{t_0} v(t)\,dt,$$

从而

$$v(t) \equiv \frac{L}{T}.$$

这与 (19.2) 矛盾.

3° 我们考虑间断的 u. 令

$$u(t)=\begin{cases}a, & t\in[0,T_0],\\ 0, & t\in(T_0,T],\end{cases}$$

其中 a,T_0 是参数, 则有解

$$v(t)=\begin{cases}at, & t\in[0,T_0],\\ aT_0, & t\in(T_0,T].\end{cases}$$

这时

$$\begin{aligned}W&=\int_0^{T_0}a^2t\,dt=\frac{a^2T_0^2}{2},\\ L&=\int_0^T v\,dt=\frac{aT_0^2}{2}+aT_0(T-T_0)=aTT_0-\frac{aT_0^2}{2}.\end{aligned}$$

当 $aT_0=\frac{L}{T}$, 且 $T_0\to 0$ 时, $W\to\frac{L^2}{2T^2}$. 对应的控制外力 $u(t)=\frac{L}{T}\delta(t)$, 其中 $\delta(t)$ 是脉冲函数.

因此脉冲力 $u=\frac{L}{T}\delta(t)$ 产生匀速 $v=\frac{L}{T}$ 实现 W 的极小值, 其解为 $W=\frac{L^2}{2T^2}$, 但脉冲力不是物理的.

4° 如果允许 u 变号, 那么可控集合变为 $U=L^\infty([0,T])$.

我们取 $u=a\left(\frac{T}{2}-t\right)$, 便得到

$$v=a\left(\frac{T}{2}t-\frac{t^2}{2}\right),\quad v(T)=0,$$

从而 $W=0$. 使得 $W=0$ 的解很多. 例如, 取

$$u(t)=\begin{cases}a, & t\in[0,T_0],\\ 0, & t\in(T_0,T-T_0),\\ -a, & t\in[T-T_0,T],\end{cases}$$

就有

$$v(t)=\begin{cases}at, & t\in[0,T_0],\\ aT_0, & t\in(T_0,T-T_0),\\ a(T-t), & t\in[T-T_0,T].\end{cases}$$

仍然是 $W=0$. 特别地, 取 $T_0=T/2$, 控制变量 u 跳跃地取正负不同的常值 $\pm a$. 这种控制称为砰砰 (bang-bang) 控制.

从这个例子可以看出, 在最优控制问题的提法中, 变量分为两类: 状态变量与控制变量.

状态变量描写系统所处的状态, 一般由向量表示 (如前例中的 x), 其变化集合为一个向量空间 $\mathbb{R}^n$ 中的子集 Y.

控制变量是人可以操作的变量 (如前例中的 u), 一切可允许的控制变量组成的集合记作 U. 它是 $L^1([0,T])$ 中的一个子集, 经常由逐段连续函数组成.

在最优控制问题中, 要给定描写系统变化规律的状态方程:

$$\dot{x} = f(t,x,u), \quad t \in (0,T), \tag{19.3}$$

要给定初态 $x(0) = x_0$ 和终态集合 $x(T) \in B$, 要通过 Lagrange 函数 $L : [0,T] \times \mathbb{R}^n \times \mathbb{R}^N \to \mathbb{R}^1$ 定义价值泛函

$$J(u) = \int_0^T L(t,x(t),u(t))dt.$$

所要确定的是 $u \in U$, 使得

$$\min_{u \in U} J(u).$$

例 19.1 (续)　在前面给出的例子中, 状态方程是

$$\begin{cases} \dot{x} = v, \\ \dot{v} = u, \end{cases} \quad t \in [0,T].$$

初态和终态集是

$$\begin{pmatrix} x(0) \\ v(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad B = \begin{pmatrix} L \\ * \end{pmatrix}.$$

可允许控制集是

$$U = \{u \in L^\infty([0,T],\mathbb{R}^1) \mid u(t) \geqslant 0\}.$$

Lagrange 函数 $L = u \cdot v$ 定义的价值泛函是

$$J(u) = W = \int_0^T uv\,dt = \int_0^T L(v,u)\,dt.$$

□

例 19.2　拿出多大的产出部分进行再投资, 可以使市场总产量极大化?

设某产品生产速度（单位时间产量）为 $q = q(t)$, 其增长的加速度 $\dot{q}$ 正比于该产品再投资的百分比 u,

$$\dot{q} = \alpha uq, \quad t \in [0,T],$$

其中 $\alpha > 0$ 是给定的, 初态也是给定的,

$$q(0) = q_0.$$

对于终态没有限制, 即 $B = \mathbb{R}^1$.

市场的总产量为

$$I = \int_0^T (1-u)q\,dt.$$

我们控制产品再投资的百分比

$$u \in U = \{v \in L^\infty(0,T) \,|\, v(t) \in [0,T]\},$$

使之极大化市场的总产量

$$\max\{I \,|\, u \in U\}. \qquad \square$$

例 19.3 (Ramsey 经济增长模型) Ramsay 早在 1928 年就提出了如下的经济增长模型.

设人均投资为 i, 人均消费为 c, 人均国民收入为 $x = i + c$.

假设 x 是人均资本总量 k 的函数 $f(k)$. 假设人口增长率为 α, 人口总量随时间增长, $L(t) = L_0 e^{\alpha t}$.

资本总量为 $K = kL$, 那么总投资为 $I = \frac{dK}{dt}$. 它们之间有如下关系:

$$i = \frac{I}{L} = \frac{1}{L}\frac{dK}{dt} = \frac{dk}{dt} + \frac{k}{L}\frac{dL}{dt} = \frac{dk}{dt} + \alpha k.$$

衡量社会消费效用的目标函数

$$W(c) = \int_0^\infty e^{-pt} u(c(t))dt,$$

其中 $p > 0$ 称为折现因子, $u(c)$ 称为消费者的效用函数. Ramsay 提出在约束

$$\frac{dk}{dt} = f(k) - \alpha k - c,\ k(0) = k_0$$

之下极大化 $W(c)$. $\square$

§19.2 Pontryagin 极大值原理

考虑下列最优控制问题.

设状态变量 $x \in W^{1,\infty}((0,T),\mathbb{R}^n)$, 控制变量 $u \in U = \{v$ 是取值于 $\mathbb{R}^N$ 的逐段连续函数 $\mid v(t) \in C\}$, 其中 $C \subset \mathbb{R}^N$ 是一个紧子集. 状态方程为 $\dot{x} = F(x,u)$, 其中 $F \in C^1(\mathbb{R}^n \times \mathbb{R}^N, \mathbb{R}^n)$, 初态 $x(0) = x_0$ 与终态 $x(T) = x_1$ 都是给定的.

又设 $L \in C(\mathbb{R}^n \times \mathbb{R}^N, \mathbb{R}^1)$ 是给定的一个 Lagrange 函数, 定义目标泛函

$$J(x,u) = \int_0^T L(x(t),u(t))dt.$$

我们有下列极大值原理.

定理 19.1 设 $u = u^*(t)$ 是上述最优控制问题的解, 对应的最优状态变量轨道是 $x = x^*(t)$, 即 $(x^*(t), u^*(t))$ 使得

$$J(x,u) = \int_0^T L(x(t),u(t))\,dt$$

达到最小值, 则存在一个逐段连续可微的函数 $\lambda = \lambda(t) \in \mathbb{R}^n$ 满足

$$\dot{\lambda} = -H_x(x^*, \lambda^*, u^*),$$

其中

$$H(x,\lambda,u) = -L(x,u) + \lambda F(x,u), \tag{19.4}$$

此外, H 沿最优轨道 $\{(x^*(t), \lambda(t), u^*(t)) \,|\, t \in (0,T)\}$ 是常数

$$E = \sup_{v \in C} H(x^*(t), \lambda(t), v).$$

我们把函数 $H(x,\lambda,u)$ 也称为 Hamilton 量, 而称 λ 为共轭变量.

证明 (这个定理的证明很长, 已超出本课程范围. 我们采用 J. Macki 和 A. Strauss 在 [MS] 给出的一个形式证明, 也参见 C. R. MacCluer [Ma].)

1° **假定**对任意初态 x^0, $\exists T_0 > 0$, $\exists u = u(x^0,t) \in U$ 极小化价值泛函. 即存在一条最优路径, $y = x(x^0,t), t \in [0,T_0]$, 满足: $\dot{y} = F(y,u), y(0) = x^0$, $y(T_0) = x_1$, 并且使得

$$\begin{aligned} I(x^0) &:= \min_{v \in U} \int_0^{T_0} L(y(t),v(t))\,dt \\ &= \int_0^{T_0} L(y(t),u(t))\,dt. \end{aligned}$$

由局部最小化原理, $(u(x^0,t), x(x^0,t))$ 在 $(t,T_0]$ 也应该是最优的, 所以有

$$I(x(x^0,t)) = \int_t^{T_0} L(x(x^0,s),u(x^0,s))\,ds.$$

2° **假设 I 是可微的**, 并且假设 u 是逐段连续的, 那么除去有限多个点外, 对任意 t 有

$$\begin{aligned}0 &= L(x(x^0,t),u(x^0,t)) + \nabla I(x(x^0,t))\cdot \dot{x}(x^0,t)\\ &= L(x(x^0,t),u(x^0,t)) + \nabla I(x(x^0,t))F(x(x^0,t),u(x^0,t)).\end{aligned}$$

当 $x^0 = x_0$ 时, 取 $u = u^*(x_0,t)$, 有 $x = x^*(x_0,t)$, 就得到

$$\begin{aligned}0 &= L(x^*(x_0,t),u^*(x_0,t)) - \lambda F(x^*(x_0,t),u^*(x_0,t))\\ &= -H(x^*(x_0,t),\lambda(t),u^*(x_0,t)),\end{aligned}$$

其中 $\lambda(t) = -\nabla I(x^*(x_0,t))$. 因为 $x^*(x_0,t)$ 是逐段可微的, 所以 λ 是逐段连续的, 并且沿最优轨道 (x^*,λ,u^*),

$$H(x^*(x_0,t),\lambda(t),u^*(x_0,t)) = 0.$$

3° 我们再来证明

$$H(x^*(t),\lambda(t),v) \leqslant 0,\ \ \forall\, v\in C,\ \forall\, t\in(0,T).$$

事实上, $\forall\, v \in C$, 方程组

$$\begin{cases}\dot{x} = F(x,v),\\ x(0) = x_0, \quad \forall t\in[0,T_0)\end{cases}$$

有局部解 $x = \tilde{x}(t)$. 一般说来, $\forall\, v\in C$, $\tilde{x}$ 不是从 x_0 到 $\tilde{x}(t)$ 的最优轨道, 我们只能有

$$I(x_0) \leqslant \int_0^t L(\tilde{x},v)\,ds + I(\tilde{x}(t)),$$

即对于足够小的 $t>0$,

$$-\frac{I(\tilde{x}(t)) - I(x_0)}{t} \leqslant \frac{1}{t}\int_0^t L(\tilde{x},v)\,ds.$$

由此推出

$$-\frac{d}{dt}I(\tilde{x}(t))|_{t=0} \leqslant L(x_0,v),$$

而左边的微分等于

$$-\nabla I(x_0)\cdot\dot{\tilde{x}}(0) = -\nabla I(x_0)\cdot F(x_0,v).$$

我们可以沿轨道 $x^*(x_0,t)$ 取任意时刻 t 的状态 $x^*(x_0,t)$ 为初值, 便得到

$$H(x^*(x_0,t),\lambda(t),v)\leqslant 0,\quad \forall v\in C,\ \forall t\in[0,T].$$

4° 最后我们证明 λ 满足共轭方程. 令

$$h(x,u)=-L(x,u)-\nabla I(x)F(x,u).$$

从前面已经证明的结论可以知道, $\forall\, t\in(0,T)$, 函数 h 除去有限多个点外, 当 $x=x^*(t),u=u^*(t)$ 时达到极大值.

假设 $I\in C^2$. 直接计算得到

$$\begin{aligned}
0&=\frac{\partial h}{\partial x_i}\\
&=-\frac{\partial L}{\partial x_i}-\frac{\partial(\nabla I\cdot F)}{\partial x_i}\\
&=-\frac{\partial L}{\partial x_i}-\frac{\partial}{\partial x_i}\left(\sum_{j=1}^N\frac{\partial I}{\partial x_j}F_j\right)\\
&=-\frac{\partial L}{\partial x_i}-\sum_{j=1}^N\frac{\partial^2 I}{\partial x_i\partial x_j}F_j-\sum_{j=1}^N\frac{\partial F_j}{\partial x_i}\frac{\partial I}{\partial x_j}\\
&=-\frac{\partial L}{\partial x_i}+\sum_{j=1}^N\frac{\partial\lambda_i}{\partial x_j}\dot{x}_j+\sum_{j=1}^N\lambda_j\frac{\partial F_j}{\partial x_i},
\end{aligned}$$

所以

$$\dot{\lambda}_i=\sum_{j=1}^N\frac{\partial\lambda_i}{\partial x_j}\dot{x}_j=\frac{\partial L}{\partial x_i}-\sum_{j=1}^N\lambda_j\frac{\partial F_j}{\partial x_i}=-\frac{\partial H}{\partial x_i}.$$

这就是

$$\dot{\lambda}=-H_x(x^*(x_0,t),\lambda(t),u^*(x_0,t)).$$

□

注 19.1 我们把 (19.4) 定义的 $H(x,\lambda,u)$ 称为 Hamilton 量, 这是因为轨道 $(x^*(t),\lambda(t),u^*(t))$ 满足下列 Hamilton 方程组:

$$\begin{cases}\dot{x}=F(t,x,u)=H_\lambda(x,\lambda,u),\\ \dot{\lambda}=-H_x(x,\lambda,u).\end{cases}$$

□

注 19.2 在这个定理中, 如果末端向量 $x(T)=x_1$ 有几个分量事先没有确定, 那么在共轭变量 $\lambda(T)$ 中相应的分量就必须是 0. □

现在我们利用极大值原理来讨论前面两个例子。

例 19.1 (再续) 倘若我们假设控制变量有约束 $0 \leqslant u(t) \leqslant a$, 其中 $a > 0$ 是一个给定的常数. 我们引入共轭变量 $\lambda(t) = (\alpha(t), \beta(t))$, 写出 Hamilton 量

$$H = -L + \lambda F = -uv + (\alpha, \beta)\begin{pmatrix} v \\ u \end{pmatrix} = \alpha v + u(\beta - v).$$

假设 $(x^*(t), v^*(t))$ 是最优轨道, 连同共轭变量的轨道 (α, β) 使得 Hamilton 量达到极大值. 因为 $v^*(t)$ 已是设定的, 所以 $u(\beta - v^*)$ 必须达到极大值. 于是有

$$u = \begin{cases} a, & \beta > v^*, \\ 0, & \beta < v^*. \end{cases}$$

由于 $\dot{v} = u, v(0) = 0$, 推出存在常数 v_0 使得

$$v^* = \begin{cases} at, & \beta > v^*, \\ v_0, & \beta < v^*. \end{cases}$$ □

我们再看共轭变量满足的方程

$$\dot{\lambda} = (\dot{\alpha}, \dot{\beta}) = -H_{(x,v)} = (0, u - \alpha).$$

从 $\lambda(t) = (\alpha(t), \beta(t))$ 推出 $\alpha = \alpha_0$ 是一个常数,

$$\beta = \begin{cases} (a - \alpha_0)t + c, & u = a, \\ -\alpha_0 t + d, & u = 0, \end{cases}$$

其中 c, d 是常数. 利用 v^* 与 β 都是连续的, 推出

$$\beta - v^* = -\alpha_0 t + c, \ d = c + v_0.$$

注意到末端 $v(T)$ 没有确定, 按注 19.2, 应当 $\beta(T) = 0$. 然而

$$-\alpha_0 t + c = 0$$

只有一个根 $T_0 = c/\alpha_0$, 所以 $c \leqslant \alpha_0 T$.

$$v^*(t) = \begin{cases} at, & t \in [0, T_0), \\ aT_0, & t \in [T_0, T]. \end{cases}$$

$$u(t) = \begin{cases} a, & t \in [0, T_0), \\ 0, & t \in [T_0, T]. \end{cases}$$

再利用终态条件 $x(T)=L$, 我们还有

$$L=\int_0^T vdt = aTT_0-\frac{aT_0^2}{2},$$

$$T_0=\begin{cases} T-\sqrt{T^2-\dfrac{2}{a}L}, & a\geqslant \dfrac{2L}{T^2},\\ \text{无解}, & a<\dfrac{2L}{T^2}. \end{cases}$$

例 19.2 (续) 回到极大化市场总产量问题.

引入共轭变量 λ, Hamilton 量是

$$H=-L+\lambda F=(1-u)q+\lambda(\alpha uq)=(1-u+\lambda\alpha u)q.$$

1° 因为 $q\geqslant 0$, 故极大化 H 等价于极大化 $-u+\lambda\alpha u=(\lambda\alpha-1)u$. 所以最优控制变量应取为

$$u(t)=\begin{cases}1, & \alpha\lambda(t)>1,\\ 0, & \alpha\lambda(t)<1.\end{cases}$$

2° 共轭变量 λ 满足方程

$$\dot{\lambda}=-H_q=-(1-u+\lambda\alpha u),$$

即 $\dot{\lambda}+\lambda\alpha u=u-1$, 故有

$$\begin{cases}\dot{\lambda}+\alpha\lambda=0, & u=1,\\ \dot{\lambda}=-1, & u=0,\end{cases}$$

得解

$$\lambda(t)=\begin{cases}ce^{-\alpha t}, & u=1,\\ -t+d, & u=0,\end{cases}$$

其中 c,d 是常数.

3° 末端条件. 因为 $q(T)$ 不确定, 所以有 $\lambda(T)=0$. 这表明 $u(T)\neq 1$, 只能是

$$u=0,\quad d=T.$$

这个结论也是合理的, 因为在生产周期之末再投资没有意义.

4° 控制变量的转换. 在 1° 中我们已经知道, 控制变量 u 的值应在 $\alpha\lambda(t)-1=0$, 即 $t_s=T-\frac{1}{\alpha}$ 时转换.

我们最终得到解

$$u(t)=\begin{cases}1, & t\in(0,t_s),\\ 0, & t\in(t_s,T),\end{cases}$$

$$q(t)=\begin{cases}q_0e^{\alpha t}, & 0<t<t_s,\\ q_0e^{\alpha t_s}, & t_s<t<T,\end{cases}$$

$$I=\int_{t_s}^{T}q_0e^{\alpha t_s}\,dt=q_0e^{\alpha t_s}(T-t_s)=\frac{q_0}{\alpha}e^{\alpha(T-\frac{1}{\alpha})}.\qquad\square$$

例 19.3 (续) 回到 Ramsey 经济增长模型.

极大化 $W(c)$ 相当于极小化 $-W(c)$. 我们引入共轭变量 λ 以及 Hamilton 量,

$$H(k,\lambda,c)=e^{-pt}u(c)+\lambda(f(k)-\alpha k-c),$$

得到最优条件

$$\frac{\partial H}{\partial c}=u'(c)e^{-pt}-\lambda=0,$$

以及

$$\begin{cases}k'(t)=f(k)-\alpha k-c,\\ \lambda'(t)=-\dfrac{\partial H}{\partial k}=-\lambda(f'(k)-\alpha),\end{cases}$$

连同初始条件 $k(0)=k_0$. 因为这是一个在半直线上的问题, 还有末端条件

$$\lim_{t\to\infty}\lambda k(t)=0.\qquad\square$$

Ramsey 早在 1928 年就用变分法来讨论这个问题, 但一直没有被人们认识到它的重要性. 直到 1965 年才由经济学家 Cass 和 Koopmans 分别加以改进. 今天它被称为 Ramsey-Cass-Koopmans 模型.

§19.3 Bang-Bang 原理

有一种控制, 其控制变量 $u\in\mathbb{R}^N$ 在全过程中每个分量最多只取两个不同的值, 我们称之为砰砰 (Bang-Bang) 控制.

给定两个矩阵 $A\in M(n\times n)$, $B\in M(n\times N)$, 设状态变量为 $x\in\mathbb{R}^n$, 控制变量为 $u\in\mathbb{R}^N$. 我们来考察下列线性系统

$$\begin{cases}\dot{x}=Ax+Bu,\\ x(0)=x^0.\end{cases}\tag{19.5}$$

定理 19.2 在线性系统 (19.5) 中, 如果存在控制变量 $u^0 \in L^\infty(0,T)$ 能使初态 $x(0)=x^0$ 变到终态 $x(T)=x^1$, 那么必有一个砰砰控制 u 实现这一目标.

这个定理实际上是泛函分析中 Krein-Milman 定理的一个推论. Krein-Milman 定理是关于局部凸拓扑线性空间中凸紧子集端点的一个结论. 它是与 Hahn-Banach 定理等价的一个命题.

设 X 是一个局部凸拓扑线性空间, E 是它的一个凸子集, 一点 $x \in E$ 称为 E 的端点, 如果不存在 $y, z \in E$, 使得 $x=(y+z)/2$. E 的端点的全体叫做 E 的端点集, 记作 $D(E)$.

Krein-Milman 定理 设 X 是一个局部凸拓扑线性空间, E 是它的一个紧凸子集, 则 E 是它的端点集 $D(E)$ 的凸闭包. □

定理 19.2 的证明 不妨设 $\|u^0(t)\|_\infty \leqslant 1$. 令

$$U=\left\{v \in L^\infty(0,T) \,|\, \|v\|_\infty \leqslant 1,\ x(T)=e^A x^0+\int_0^t e^{A(t-s)}Bv(s)ds\right\}.$$

显然 U 是一个 *弱闭凸集, 并且 $u^0 \in U$.

按 Banach-Alaoglu 定理, $L^\infty(0,T)$ 中的单位球是 *弱紧的, 从而 U 也是 *弱紧集. 按 Krein-Milman 定理, 它有端点.

以下我们来证明: 如果 $u \in U$ 是 U 的一个端点, 那么 u 必是一个砰砰控制. 为简单计, 仅设 $N=1$. 任取 $\varepsilon \in (0,1)$, 考察集合

$$S=\{t \in (0,T) \,|\, u(t) \in (-1+\varepsilon, 1-\varepsilon)\}.$$

注意到线性系统的解可以表示为

$$x(t)=e^A x^0+\int_0^t e^{A(t-s)}Bu(s)ds,$$

我们取一个函数 $v \in L^\infty(0,T)$, 满足

$$\int_S e^{-As}Bv(s)ds=0.$$

令 v 在 S 外为 0, 适当调节倍数使得 $\|v\|_\infty=1$. 于是 $u \pm \varepsilon v \in U$, 并且 $u=[(u+\varepsilon v)+(u-\varepsilon v)]/2$. 这个矛盾表明 $S=\varnothing$. 而 $\varepsilon>0$ 是任意的, 所以 u 是一个砰砰控制. □

第二十讲　有界变差函数与图像恢复

在一些现代应用问题中 (如图像恢复、相变等问题), 变分问题出现的解是不连续的, 而且这种不连续性表现为有明显断裂的可求长曲线 (或可度量的高维曲面). 这表明所求的解不在相关的 Sobolev 空间中, 人们转向在有界变差函数空间中求解.

从 20 世纪中期开始, De Giorgi, Reifenberg, Federer, Fleming 和 Almgren 等人在研究极小曲面的 Plateau 问题时, 就采用测度论的方法. 有界变差函数与测度有天然的联系, 与此相关的理论框架已经系统地发展成为几何测度论. 但这些内容已远远超出本课程的范围, 我们只能通过一些具体应用的例子来了解其大概.

我们先回忆一个变量的有界变差函数的定义和主要性质, 然后再看怎样把它推广到多个变量的情形.

§20.1　一元有界变差函数的回顾

定义 20.1　函数 $u: J=(a,b)\to\mathbb{R}^1$ 称为是有界变差的 (bounded variation), 简作 BV, 如果 $\exists M>0$ 使得对于 J 的任意一个分划 $\pi: a<t_1<\cdots<t_{n+1}<b$, 都有

$$S_\pi(u)=\sum_{i=1}^n |u(t_{i+1})-u(t_i)|\leqslant M.$$

我们把

$$V_a^b(u)=\sup_\pi S_\pi(u)$$

称为 u 的全变差 (total variation).

J 上全体 BV 函数组成的集合记作 BV(J), 它是一个线性空间. 下列简单性质都是直接可以推出来的:

(1) J 上的有界单调函数是 BV 函数.

(2) $\mathrm{Lip}(J,\mathbb{R}^1) \subset \mathrm{BV}(J)$.

(3) $\forall u \in \mathrm{BV}(J)$, u 可以分解为两个递增函数之差:

$$u(x) = \frac{1}{2}(T_u(x) + u(x)) - \frac{1}{2}(T_u(x) - u(x)),$$

其中

$$T_u(x) = \sup\left\{ \sum_{i=1}^{m} \left|u(t_{i+1}) - u(t_i)\right| \sigma : a < t_1 < \cdots < t_{m+1} < x \text{ 是任意一个分划} \right\}.$$

(4) $\forall u \in \mathrm{BV}(J)$, 跳跃点集 $D_u = \{x \in J \mid u(x-0) \neq u(x+0)\}$ 至多是可数集.

BV 函数的值在跳跃点上是不确定的. 要避开由此带来的麻烦, 我们将其规范为

$$u(x) = u(x-0), \quad \forall x \in J.$$

为了突出 BV 函数在 J 上的变差部分, 我们再规范 $u(a+0) = 0$. 规范后的 BV 函数全体还是一个线性空间, 记作 NBV(J). 对它引入范数

$$\|u\| = V_a^b(u).$$

NBV(J) 是一个 Banach 空间.

下面我们来建立 NBV 函数与完全可加 Borel 可测集函数之间的一一对应关系.

任意给定一个单调递增的规范有界变差函数 $u : J \to \mathbb{R}^1$, 定义

$$\mu_u((a,x)) = u(x).$$

我们说 μ_u 可以扩张为 J 上的一个测度, 即完全可加的非负集函数. 事实上, $\forall [\alpha,\beta) \subset J$, 令

$$\mu_u([\alpha,\beta)) = u(\beta) - u(\alpha),$$

便有 $\mu_u(\{\alpha\}) = u(\alpha+0) - u(\alpha)$. 从而 $\forall$ Borel 子集 $E \subset J$, 有

$$\mu_u(E) = \mathcal{L}^1\left(\bigcup_{x\in E} [u(x), u(x+0)]\right),$$

其中 $\mathcal{L}^1$ 是 $\mathbb{R}^1$ 上的 Lebesgue 测度.

反之, 如果给定了 J 上的一个 Borel 测度 μ, 那么函数 $u(x) = \mu((a,x))$ 是单调递增的并且是左连续的, 还满足 $u(a+0) = 0$.

再由简单性质 (3), $\forall u \in \mathrm{BV}(J)$ 存在两个单调递增的左连续函数 u_1 与 u_2, 使得 $u = u_1 - u_2$. 于是按上面的结论, 对应着两个 Borel 测度 $\mu_i = \mu_{u_i}$, $i = 1, 2$. 若令

$$\mu(E) = \mu_1(E) - \mu_2(E),$$

则 μ 是一个完全可加的 Borel 可测集函数, 满足

$$\mu((a,x)) = \mu_1((a,x)) - \mu_2((a,x)) = u(x), \quad \forall x \in J.$$

反之, 对任意完全可加的 Borel 可测集函数 μ, 通过

$$u(x) = \mu((a,x)) \tag{20.1}$$

定义出唯一的一个规范后的有界变差函数.

这样一来, 我们对任意完全可加 Borel 可测集函数 μ 以及任意 Borel 可测集 $E \in \mathcal{B}$, 也可以定义 "全变差":

$$|\mu|(E) = \sup\left\{ \sum_{i=1}^{\infty} |\mu(B_i)| \,\middle|\, B_i \in \mathcal{B}, B_i \cap B_j = \varnothing,\ i \neq j, \bigcup_{i=1}^{\infty} B_i = E \right\}.$$

不难验证

(1) $|\mu|$ 是一个 Borel 测度,

(2) $|\mu(E)| \leqslant |\mu|(E), \forall E \in \mathcal{B}$,

(3) $|\mu|(J) = V_a^b(u)$.

我们用 $\mathcal{M}(J, \mathbb{R}^1)$ 表示 J 上完全可加 Borel 可测集函数全体组成的空间. 在 $\mathcal{M}(J, \mathbb{R}^1)$ 上定义模

$$\|\mu\|_{\mathcal{M}} = |\mu|(J),$$

它也成为一个 Banach 空间.

以上建立起来的一一对应 $u \mapsto \mu$ 是 Banach 空间 $\mathrm{NBV}(J)$ 与 $\mathcal{M}(J, \mathbb{R}^1)$ 间的同构.

现在我们回顾一下连续函数空间上连续线性泛函表示的 Riesz 定理.

$$C_0(J)^* \cong \mathcal{M}(J, \mathbb{R}^1),$$

其中 $C_0(J) = \{\varphi \in C(\bar{J}) \mid \varphi(a) = \varphi(b) = 0\}$.

因为 $C_0(J)$ 是可分的, 所以 $\mathcal{M}(J, \mathbb{R}^1)$ 是可分 Banach 空间的共轭空间.

利用 Lebesgue-Nikodym 分解定理, 我们知道任意一个 BV 函数可以有如下分解:

$$u(x) = v(x) + r(x) + s(x),$$

其中 $v(x)$ 是绝对连续部分, $r(x)$ 为 Cantor 部分, 而 $s(x)$ 为跳跃部分. 特别地, s 是一个跳跃函数, 而 r 是一个有几乎处处为零的导数, 但本身不等于常数的连续有界变差函数.

注意到BV 函数 u 是有界可测函数, 从而本身属于 $L^1(J)$. 作为 L^1 函数, 它有 (在广义函数意义下的) 广义导数 Du,

$$\langle Du, \varphi \rangle = -\int_J u \cdot \varphi' \, dx, \quad \forall \varphi \in C_0^\infty(J). \tag{20.2}$$

请读者特别注意, 与 Sobolev 空间不同, 对于有界变差函数, Du 与几乎处处导数 u' 并不相同! 我们知道只有当 $r = s = 0$ 时, $u \in AC(J)$, 才有 $Du = u'$.

如果 $s \neq 0$, 那么 Du 可以是测度. 例如, $J = (-1, 1)$,

$$u(x) = \begin{cases} 0, & x \leqslant 0, \\ 1, & x > 0, \end{cases}$$

这时 $Du = \delta(x)$. 即, 若令 $\mu = Du$, 则有 $\mu(\{0\}) = 1, \mu(B) = 0, \forall B \in \mathcal{B}, 0 \notin B$.

现在我们再把有界变差函数 u 与它对应的测度 μ 之间的关系更明确地表达出来. 注意到 $\mathcal{M}(J, \mathbb{R}^1)$ 作为 $C_0(J)^*$ 有如下的对偶关系,

$$\langle \mu, \varphi \rangle = \int_J \varphi \, d\mu, \quad \forall (\mu, \varphi) \in \mathcal{M}(J, \mathbb{R}^1) \times C_0(J),$$

应有

$$\|\mu\|_{\mathcal{M}} = \sup \left\{ \int_J \varphi \, d\mu \middle| \|\varphi\|_{C_0} \leqslant 1 \right\}.$$

按照 $\mathcal{M}(J, \mathbb{R}^1)$ 与 $\mathrm{NBV}(J)$ 的一一对应, 还有

$$\langle \mu, \varphi \rangle = \int_J \varphi \, du(x) = -\int_J u(x) \varphi'(x) \, dx.$$

联系到 (20.2), 即得

$$\boxed{Du = \mu.}$$

多元函数虽没有左、右端点, 无从用来规范函数值, 但上述公式对于把有界变差函数推广到多个变量有启发.

定义 20.2 在 BV(J) 上定义模:

$$\|u\|_{BV} = \|u\|_{L^1} + \|Du\|_{\mathcal{M}}.$$

不难验证, BV(J) 是一个 Banach 空间.

§20.2 多元有界变差函数

设 $\Omega \subset \mathbb{R}^n$ 是一个有界开集. 用 $C_0^1(\Omega, \mathbb{R}^n)$ 表示具有紧支集的 $\mathbb{R}^n$ 值连续可微函数空间. 仿照一元函数情形, 引入如下定义.

定义 20.3 $u \in L^1(\Omega)$ 称为是有界变差的, 如果

$$\|Du\|(\Omega) = \sup\left\{ \int_\Omega u \mathrm{div}\varphi \, dx \,\middle|\, \varphi \in C_0^1(\Omega, \mathbb{R}^n), \|\varphi\|_\infty \leqslant 1 \right\} < \infty.$$

我们把 $\|Du\|(\Omega)$ 称为 u 的全变差.

Ω 上一切 BV 函数构成的空间记作 BV(Ω), 引入模

$$\|u\|_{\mathrm{BV}} = \|u\|_{L^1} + \|Du\|(\Omega).$$

和一元函数一样, 对任意 BV 函数 u, 我们考察它的广义函数梯度 Du:

$$\langle Du, \varphi \rangle = \langle u, \mathrm{div}\varphi \rangle := \int_\Omega u \mathrm{div}\varphi dx \quad (\forall\, \varphi \in C_0^\infty(\Omega))$$

与测度间的关系.

可以说由 u 的全变差有界可以认定 Du 是一个向量值的 Radon 测度.

这是因为 $\forall\, \varphi_0 \in C_0(\Omega, \mathbb{R}^n)$, 在 $\bar{\Omega}$ 上一切边值为零的连续函数组成的 Banach 空间, $\exists \{\varphi_k\} \subset C_0^1(\Omega, \mathbb{R}^n)$, 使得 φ_k (在 $\bar{\Omega}$ 上) 一致收敛于 φ_0, 且 $|\varphi_k|_\infty \leqslant |\varphi_0|_\infty, \forall\, k$. 定义线性泛函

$$L(\varphi) = \int_\Omega u \, \mathrm{div} \, \varphi \, dx, \quad \forall\, \varphi \in C_0^1(\Omega, \mathbb{R}^n),$$

它满足

$$|L(\varphi)| \leqslant C\|\varphi\|_\infty, \quad \forall\, \varphi \in C_0^1(\Omega, \mathbb{R}^n).$$

如今令

$$\overline{L}(\varphi_0) \doteq \lim_{k\to\infty} L(\varphi_k),$$

则极限 $\overline{L}(\varphi_0)$ 存在, 而且不依赖于 φ_k 的特殊选择. 这表明 L 有唯一的线性连续扩张:

$$\overline{L}: C_0(\Omega, \mathbb{R}^n) \to \mathbb{R}^1,$$

即, $\bar{L}$ 是 $C_0(\Omega, \mathbb{R}^n)$ 上的线性连续泛函. 利用Riesz表示定理, 存在Radon测度 μ 与一个 μ-可测函数 σ, 使得

$$|\sigma(x)| = 1, \quad \mu \text{ a.e. 于 } \Omega,$$

以及

$$\int_\Omega u \operatorname{div} \varphi \, dx = -\int \sigma \cdot \varphi \, d\mu, \quad \forall \varphi \in C_0^1(\Omega, \mathbb{R}^n).$$

这表明

$$\boxed{Du = \sigma \, \mathrm{d}\mu.}$$

今后我们用 $\mathcal{M}(\Omega, \mathbb{R}^{n+1})$ 记 Ω 上的 $\mathbb{R}^{n+1}$ 值Radon 测度组成的空间, 并定义模

$$\|\mu\|_{\mathcal{M}} = \sup\left\{ \int_\Omega \varphi \, d\mu \middle| \|\varphi\|_{C_0(\Omega, \mathbb{R}^{n+1})} \leqslant 1 \right\},$$

使之成为一个Banach空间.

例 20.1 设 $u \in W^{1,1}(\Omega)$, 则 $u \in \mathrm{BV}(\Omega)$, 且 $\|u\|_{W^{1,1}} = \|u\|_{\mathrm{BV}}$. 事实上, 一方面,

$$\left| \int_\Omega u \operatorname{div} \phi dx \right| = \left| \int Du \cdot \phi dx \right| \leqslant \|Du\|_1, \quad \forall \phi \in C_0^1(\Omega, \mathbb{R}^n), \|\phi\|_\infty \leqslant 1.$$

另一方面, $\forall \varepsilon > 0, \exists \varphi_\varepsilon \in C_0^1(\Omega, \mathbb{R}^n), \|\varphi_\varepsilon\|_\infty \leqslant 1$, 满足

$$\int_\Omega |Du| dx \leqslant \int_\Omega Du \cdot \varphi_\varepsilon \, dx + \varepsilon = \int_\Omega u \operatorname{div} \varphi_\varepsilon \, dx + \varepsilon \leqslant \|Du\|(\Omega) + \varepsilon,$$

即得 $\|Du\|(\Omega) = \|Du\|_1$. □

例 20.2 设 $S \subset \mathbb{R}^n$ 是一个 C^∞ $(n-1)$ 维紧超曲面. 按 $\mathbb{R}^n$ 上的度量导出 S 上的度量, 其 $(n-1)$ 维 Hausdorff 测度记作 $\mathcal{H}^{n-1}(S)$.

又设 Ω 是由 S 围成的区域, 而 χ_Ω 为其特征函数, 则 $\chi_\Omega \in \mathrm{BV}(\mathbb{R}^n)$, 并且

$$\|D\chi_\Omega\|(\mathbb{R}^n) = \mathcal{H}^{n-1}(S).$$

证明 1° $\forall \phi \in C_0^1(\mathbb{R}^n, \mathbb{R}^n)$, 由 Gauss 公式,

$$\int_{\mathbb{R}^n} \chi_\Omega \cdot \operatorname{div} \phi \, dx = \int_\Omega \operatorname{div} \phi \, dx = \int_S n(x) \cdot \phi(x) \, d\mathcal{H}^{n-1},$$

其中 $n(x)$ 是单位法向量. 因此,

$$\|D\chi_\Omega\|(\mathbb{R}^n) \leqslant \mathcal{H}^{n-1}(S).$$

2° 另一方面, 可以用单位分解把 $n(x)$ 延拓到 $\mathbb{R}^n$ 上成为一个 C^∞ 向量场 V, 满足 $\|V(x)\| \leqslant 1, \forall x \in \mathbb{R}^n$.

$\forall \rho \in C_0^\infty(\mathbb{R}^n, \mathbb{R}^1), |\rho(x)| \leqslant 1, \forall x \in \mathbb{R}^n$, 令 $\phi = \rho V$. 我们有

$$\int_{\mathbb{R}^n} \chi_\Omega \operatorname{div} \phi \, dx = \int_S \rho \, d\mathcal{H}^{n-1}.$$

因此,

$$\begin{aligned}
\|D\chi_\Omega\|(\mathbb{R}^n) &\geqslant \sup\left\{ \int_\Omega \chi_\Omega \operatorname{div} \phi \, dx \middle| \phi \in C_0^\infty(\mathbb{R}^n, \mathbb{R}^n), \|\phi\|_\infty \leqslant 1 \right\} \\
&\geqslant \sup\left\{ \int_S \rho \, \mathrm{d}\mathcal{H}^{n-1} \middle| \rho \in C_0^\infty(\mathbb{R}^n, \mathbb{R}^1), |\rho(x)| \leqslant 1, \forall x \in \mathbb{R}^n \right\} \\
&= \mathcal{H}^{n-1}(S).
\end{aligned}$$

□

从这两个例子可以看出 $W^{1,1}(\Omega) \subset \mathrm{BV}(\Omega)$, 但不相等.

$\mathrm{BV}(\Omega)$ **有下列简单性质**.

(1) (下半连续性) 若 $\{u_j\} \subset \mathrm{BV}(\Omega)$, 且 $u_j \to u \ (L^1(\Omega))$, 则 $\forall$ 开集 $U \subset \Omega$,

$$\|Du\|(U) \leqslant \underline{\lim} \|Du_j\|(U).$$

进一步, 由 $\sup\{\|Du_j\|(\Omega) \mid j = 1, 2, \cdots\} < \infty$, 可以推出 $\|Du\|(\Omega) < \infty$.

证明 $\forall \phi \in C_0^1(\Omega, \mathbb{R}^n), \|\phi\|_\infty \leqslant 1$, 我们有

$$\int_\Omega u \operatorname{div} \phi \, dx = \lim_{j \to \infty} \int_\Omega u_j \operatorname{div} \phi \, dx \leqslant \underline{\lim} \|Du_j\|(\Omega).$$

这就是结论. □

(2) $\mathrm{BV}(\Omega)$ 是完备的.

证明 设 $\{u_j\} \subset \mathrm{BV}(\Omega)$ 是一个 Cauchy 序列. 由定义, 它们是 $L^1(\Omega)$ 中的 Cauchy 序列. 于是有 $u_j \to u \in L^1(\Omega)$. 应用性质 (1), $u \in \mathrm{BV}(\Omega)$. 下面证明:

$$\|D(u_j - u)\|(\Omega) \to 0.$$

事实上, 再应用性质 (1), $\forall \varepsilon > 0, \exists j_0 > 0$,

$$\|D(u_j - u)\|(\Omega) \leqslant \varliminf_{k\to\infty} \|D(u_j - u_k)\|(\Omega) < \varepsilon, \quad j \geqslant j_0. \qquad \square$$

(3) (逼近) $\forall u \in \mathrm{BV}(\Omega), \exists \{u_j\} \subset \mathrm{BV}(\Omega) \cap C^\infty(\Omega)$, 使得

(a) $u_j \to u$ 于 $L^1(\Omega)$,

(b) $\|Du_j\| \to \|Du\|$ 在 Radon 测度意义下. 特别地, $\|Du_j\|(\Omega) \to \|Du\|(\Omega)$.

证明参看 [AFP] pp.122–123.

(4) (紧性) 设 $\{u_j\} \subset \mathrm{BV}(\Omega)$ 满足 $\|u_j\|_{\mathrm{BV}} \leqslant M$, 则存在子列 $\{u_{n_j}\}$ 及 u 使得 $u_{n_j} \to u(L^1)$, $\|Du\|(\Omega) \leqslant M$.

证明 由性质 (3), $\exists \{v_j\} \subset \mathrm{BV}(\Omega) \cap C^\infty(\Omega)$, 使得 $\|v_j - u_j\|_1 \leqslant 1/j$, $\|v_j\|_{\mathrm{BV}} \leqslant \|u_j\|_{\mathrm{BV}} + \frac{1}{j}$.

因为 $\|v_j\|_{W^{1,1}} = \|v_j\|_{\mathrm{BV}}$, 所以 $\{v_j\}$ 是 $W^{1,1}$有界的. 再由 Rellich-Kondrachov 定理, 存在子列 n_j 使得 $v_{n_j} \to u$ $(L^1(\Omega))$, 推出 $u_{n_j} \to u$ $(L^1(\Omega))$. 再应用性质 (1), 得

$$\|Du\|(\Omega) \leqslant \varliminf \|Dv_j\|(\Omega) \leqslant \varliminf \|Du_j\|(\Omega) \leqslant M. \qquad \square$$

此外, BV 函数也有 Poincaré 不等式.

(5) Poincaré 不等式. 设 $\Omega \subset \mathbb{R}^n$ 是一个有界连通可延拓的区域, 则有不等式

$$\left\| u - \frac{1}{|\Omega|} \int_\Omega u\, dx \right\|_p \leqslant C_p \|Du\|(\Omega), \quad \forall u \in \mathrm{BV}(\Omega), 1 \leqslant p \leqslant \frac{n}{n-1}.$$

参看 [AFP] p.153.

以下我们来考察 $\mathrm{BV}(\Omega)$ 的 *弱收敛. 记 $X = C_0(\Omega, \mathbb{R}^{n+1})$. 令

$$E = \{\phi = (\varphi_0, \varphi_1, \cdots, \varphi_n) \in C_0^\infty(\Omega, \mathbb{R}^{n+1}) \mid \operatorname{div} \hat{\varphi} = \varphi_0, \hat{\varphi} = (\varphi_1, \cdots, \varphi_n)\},$$

Y 是 E 在 X 中之闭包. 因此 Y 是 X 的一个闭线性子空间.

考虑线性映射

$$\begin{aligned} T: \mathrm{BV}(\Omega) &\to \mathcal{M}(\Omega, \mathbb{R}^{n+1}), \\ u \quad &\mapsto (u\mathcal{L}^n, D_1 u, \cdots, D_n u), \end{aligned}$$

其中 $\mathcal{L}^n$ 是 n-Lebesgue 测度, 而 $D_i u$ 则是 Radon 向量测度 Du 的第 i 个分量, 在广义函数意义下,

$$\langle D_i u, \varphi \rangle = -\langle u, \partial_{x_i} \varphi \rangle, \quad \forall \varphi \in C_0^\infty(\Omega),\ 1 \leqslant i \leqslant n.$$

事实上, $Tu \in (X/Y)^*$. 这是因为 $\forall \phi \in E$,

$$\langle Tu, \phi\rangle = \langle u\mathcal{L}^n, \varphi_0\rangle + \langle Du, \hat{\varphi}\rangle = \langle u\mathcal{L}^n, \varphi_0 - \operatorname{div}\hat{\varphi}\rangle = 0.$$

即 $Tu|_Y = 0$, 亦即 $Tu \in (X/Y)^*$.

T 还是满射, 即 $\forall (\mu_0, \mu_1, \cdots, \mu_n) \in \mathcal{M}(\Omega, \mathbb{R}^{n+1})$, 如果

$$\langle \mu_0, \varphi_0\rangle + \sum_{i=1}^{n} \langle \mu_i, \varphi_i\rangle = 0, \quad \forall (\varphi_0, \cdots, \varphi_n) \in E,$$

那么

$$\langle \mu_i, \varphi\rangle = -\langle \mu_0, \partial_i\varphi\rangle, \quad i = 1, 2, \cdots, n, \forall \varphi \in C_0^1(\Omega).$$

利用测度论的技巧, 不难验证 μ_0 实际上是绝对连续的, 从而存在 $u \in L^1(\Omega)$, 使得 $\mu_0 = u\mathcal{L}^n$.

注意到

$$\|Tu\|_{\mathcal{M}(\Omega,\mathbb{R}^{n+1})} = \sup\left\{ \int_\Omega (u \cdot \varphi_0 + Du \cdot \hat{\varphi})\, dx \,\middle|\, \|\phi\|_{C_0(\Omega,\mathbb{R}^n+1)} \leqslant 1 \right\},$$

可见

$$\|u\|_{\mathrm{BV}} = \|u\|_{L^1} + \|Du\|(\Omega) \leqslant 2\|Tu\| \leqslant 2\|u\|_{\mathrm{BV}}.$$

这样一来, 我们便把 $\mathrm{BV}(\Omega)$ 看成一个可分 Banach 空间 (X/Y) 的共轭空间. 由 Banach-Alaoglu 定理, 其有界集是 *弱列紧的.

现在我们再来了解 $\mathrm{BV}(\Omega)$ 的 *弱拓扑.

定理 20.1

$$u_j \xrightarrow{\mathrm{BV}\ ^*\text{弱}} u \Longleftrightarrow \begin{cases} u_j \to u,\ L^1, \\ u_j \text{ 在 BV 中有界}. \end{cases}$$

证明

"$\Rightarrow$". 由 Banach-Steinhaus 定理直接推出 $\{u_j\}$ 在 BV 中是有界的.

再利用性质 (4), u_j 在 L^1 中收敛到 u.

"$\Leftarrow$". 由 Banach-Alaoglu 定理, $\{u_j\}$ 中有子列 $u_{n_j}^*$ 弱收敛. 剩下来只需证明 $\{Du_j\}$ 的所有可能的 *弱极限点是同一个 Du.

事实上若存在子列 $\{u_{n_j}\}$ 使得 $\lim Du_{n_j} = \mu$, 则

$$\int_\Omega u_{n_j} \operatorname{div}\varphi\, dx = -\int_\Omega Du_{n_j} \cdot \varphi\, dx,$$

推得

$$\int_{\Omega} u \operatorname{div} \varphi \, dx = -\langle \mu, \varphi \rangle.$$

此即 $\mu = Du$. □

多元 BV 函数 u 在广义函数意义下的广义梯度 Du 也有 Lebesgue 分解:

$$Du = D^a u + D^j u + D^c u,$$

其中

$$D^a u = \nabla u \mathcal{L}^n \in L^1(\Omega, \mathbb{R}^n)$$

是绝对连续部分, $D^j u$ 与 $D^c u$ 分别称为跳跃部分与 Cantor 部分, 其中 $\mathcal{L}^n$ 是 $\mathbb{R}^n$ 上的 Lebesgue 测度.

它们是这样定义的: 记 $B_r(x)$ 为 $\mathbb{R}^n$ 中以 x 为中心, $r > 0$ 为半径之球. 定义两个函数 (图 20.1)

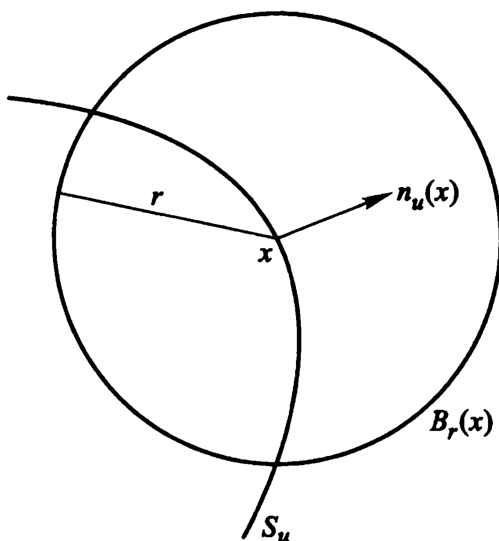


图 20.1

$$u^+(x) = \inf \left\{ t \in [-\infty, +\infty] \,\middle|\, \lim_{r \to 0} \frac{\mathcal{L}^n \{ y \in B_r(x) \,|\, u(y) > t \}}{r^n} = 0 \right\},$$
$$u^-(x) = \sup \left\{ t \in [-\infty, +\infty] \,\middle|\, \lim_{r \to 0} \frac{\mathcal{L}^n \{ y \in B_r(x) \,|\, u(y) < t \}}{r^n} = 0 \right\}.$$

一点 $x \in \Omega$ 称为是 u 的 Lebesgue 点, 如果

$$\lim_{r \to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |u(y) - u(x)| \, dy = 0.$$

这时有

$$u(x)=\lim_{r\to 0}\frac{1}{|B_r(x)|}\int_{B_r(x)}u(y)\,dy,$$

以及

$$u(x)=u^+(x)=u^-(x).$$

我们把集合

$$S_u=\{x\in\Omega\mid u^+(x)>u^-(x)\}$$

称为 u 的跳跃集. 可以证明跳跃集 S_u 是一个可数可求长的集合.

对于 $(n-1)$ 维 Hausdorff 测度 $\mathcal{H}^{n-1}$, 几乎所有的 $x\in S_u$ 都可以定义法向 $n_u(x)$. 然后定义

$$D^j u=(u^+-u^-)n_u\cdot\mathcal{H}^{n-1}|_{S_u},$$

以及

$$D^c u=(Du-D^a u)\lfloor(\Omega\backslash S_u).$$

参看 [AFP] pp. 187–189.

§20.3 松弛函数

在一定增长性的限制下, 我们知道泛函 I 的弱序列下半连续性与 Lagrange 函数 L 对 p 的拟凸性是等价的. 在第十三讲有一个 Bolza 的反例 (例 13.3), 即 $L(u,p)=u^2+(p^2-1)^2$. 它对 p 当然不是凸的, $I(u)$ 在 $W^{1,4}(0,1)$ 上有极小化序列 u_n 弱收敛于 0, 相应的泛函值 $I(u_n)\to 0$, 但 $I(0)=1$!

然而从求泛函极值的角度看, 在这个例子中, 泛函 I 在一点 $u=0$ 处的值并不那么重要, 倒是极小化序列反映了泛函 I 可能达到多小的程度.

为了突现泛函的这方面特征, 我们引入松弛函数的概念.

定义 20.4 设 (X,τ) 是一个拓扑空间, $f:X\to\bar{\mathbb{R}}=\mathbb{R}^1\cup\{+\infty\}$, 称 $Rf(x)=\sup\{\varphi(x)\mid\varphi:X\to\bar{\mathbb{R}}$ 下半连续, $\varphi(y)\leqslant f(y),\ \forall y\in X\}$ 为 f 的松弛函数.

由定义立得如下结论.

(1) 若 f 是下半连续的, 则 $Rf=f$.

证明 事实上, 取 $\varphi=f$, 可见 $f\leqslant Rf$. 另一方面, 由定义 $Rf\leqslant f$. □

(2) $Rf(x)$ 是下半连续的. 这是因为下半连续函数族的上确界函数是下半连续的.

由此可见, f 的松弛函数 Rf 是小于等于 f 的诸多下半连续函数中最大的.

(3) 设有下界的函数 $f: X \to \bar{\mathbb{R}}$ 有极小化序列 $\{x_j\}$, 而 x_0 是 $\{x_j\}$ 的一个极限点, 则 x_0 是 Rf 的极小点, 而且 $Rf(x_0) = \inf f$.

证明 不妨设 $c = \inf f = \lim f(x_j)$, 那么一方面,

$$\begin{aligned} Rf(x_0) &\leqslant \underline{\lim} Rf(x_j) && \text{(由 (2))} \\ &\leqslant \underline{\lim} f(x_j) = c && \text{(由定义)}. \end{aligned}$$

另一方面, $c \leqslant f(x)$, $\forall x \in X$. 作为常值函数 c 是弱下半连续的, 从而 $c \leqslant Rf(x)$, $\forall x \in X$. 这表明 $Rf(x_0) = c$, 并且 x_0 是 Rf 的极小点. □

定理 20.2 设 X 是一个可分的 Banach 空间的共轭空间. 若有下界的函数 $f: X \to \mathbb{R}^1$ 是强制的, 则按 * 弱拓扑定义的松弛函数 Rf 有极小点, 且

$$\min_X Rf = \inf_X f.$$

证明 取 f 的极小化序列 $\{x_j\}: \lim f(x_j) = \inf_X f$, $\{x_j\}$ 是 X 中的有界列, 从而有 * 弱收敛子列 $x_{n_j} \xrightarrow{*w} x_0$. 按 * 弱拓扑定义的松弛函数 Rf 是 * 弱下半连续的, 由性质 (3), x_0 是 Rf 的极小点,

$$\min_X Rf = Rf(x_0) = \inf_X f.$$ □

现在回到变分问题. 给定一个 Lagrange 函数 $L: \mathbb{R}^n \to \mathbb{R}^1$ 是连续的, 满足增长条件

$$c_0|p|^q \leqslant L(p) \leqslant c_1|p|^q + c_2, \quad 1 < q < \infty.$$

定义泛函

$$I(u) = \int_\Omega L(\nabla u(x))\, dx, \quad u \in W^{1,q}(\Omega), 1 < q < \infty.$$

因为此时

$$L \text{ 弱下半连续} \Leftrightarrow L \text{ 拟凸} \Leftrightarrow L \text{ 凸},$$

所以我们有

$$RI(u) = \int_\Omega \operatorname{conv}(L)(\nabla u(x))\, dx,$$

其中 $\operatorname{conv}(L) = \sup\{\varphi \mid \varphi(x) \leqslant L(x), \forall x \in \Omega, \varphi \text{ 是凸的}\}$ 表示 L 的凸包, 即不超过 L 的凸函数的上确界. □

例 20.1 (双井) 设 $L(t,u,p)=(1-p^2)^2$ (见图 20.2),

$$I(u)=\int_{-1}^{1}L(t,u(t),\dot{u}(t))dt.$$

令

$$\tilde{L}(t,u,p)=\begin{cases}(1-p^2)^2, & |p|>1,\\ 0, & |p|\leqslant 1.\end{cases}$$

它是 L 的凸包, 则

$$RI(u)=\int_{-1}^{1}\tilde{L}(t,u(t),\dot{u}(t))dt.$$

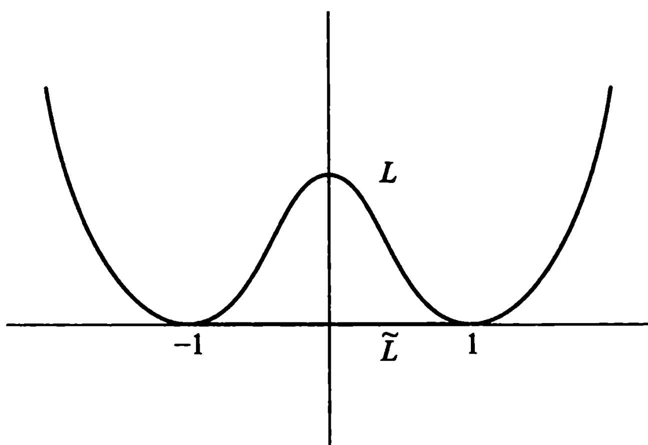


图 20.2

在近代变分学中有大量文献研究给定泛函的松弛泛函的具体表现形式, 参看 [Bu] G. Buttazzo (1989). □

§20.4 图像恢复与 Rudin-Osher-Fatemi 模型

光学仪器所成的像总有失真之处. 除了噪音而外, 仪器本身也有点扩散函数影响成像效果.

设 Ω 是一个区域, 真实景像用 Ω 上的函数表示, $u:\Omega\subset\mathbb{R}^2\to\mathbb{R}$, 用光学仪器观察到的像是

$$u_d(x)=Ku(x)+n(x)=\int k(x-y)u(y)dy+n(x),$$

其中 K 是光学仪器自身的点扩散函数 $k(x)$ 导出的卷积算子. 它可以看成是 $L^2(\Omega)$ 上的一个有界线性算子, 而 n 是噪音. 所谓图像恢复就是由已知 u_d, 反过来求 u.

最自然的方法是求解最小二乘问题

$$\inf \int_\Omega |u_d - Ku|^2\, dx, \tag{20.1}$$

它的E-L方程是

$$K^* u_d - K^* K u = 0, \tag{20.2}$$

其中 K^* 是 K 的共轭算子. 因为 K^*K 并没有有界逆, 方程 (20.2) 一般来说是*非适定的*.

Tikhonov 和 Arsenin 于 1977 年提出了一种*正则化方法*, 在 (20.1) 的目标泛函上添加一个能量项:

$$T_\lambda(u) = \int_\Omega |u_d - Ku|^2\, dx + \lambda \int_\Omega |\nabla u|^2\, dx, \quad \lambda > 0 \text{ 为参数}.$$

这样做的好处是既能克服 (20.2) 的非适定性, 又能使解 u 变得光滑.

如果我们在空间 $H^1(\Omega)$ 上求这个泛函的极小值, 那么它的E-L方程就是

$$K^* u_d - K^* K u + \lambda \Delta u = 0,$$

连同Neumann边值条件

$$\frac{\partial u}{\partial n}\Big|_{\partial\Omega} = 0.$$

椭圆方程解的正则性保证了解 u 是光滑的. 这样自然也就把一些噪音过滤掉了. 但这种各向同性的光滑化也有问题, 因为它把图像中物体的边缘也弄得模糊了!

人们试图用梯度 ∇u 的 L^p 模取代 L^2 模, 并减小 p, 以尽量保持物体的边缘. Rudin-Osher-Fatemi提出用 L^1 模, 即用 $\int_\Omega |\nabla u|\, dx$ 取代 $\int_\Omega |\nabla u|^2\, dx$.

然而空间 $W^{1,1}(\Omega)$ 既不是自反的, 也不是什么空间的共轭空间, 在它上面难以判断泛函极小值的存在性.

我们前面说过, 一方面, $\mathrm{BV}(\Omega)$ 是一个可分 Banach 空间的共轭空间; 另一方面, BV 解 u 可能出现间断, 间断集可以是可求长的曲线. 如果在 $\mathrm{BV}(\Omega)$ 空间中求解, 那么不但在存在性方面有便利, 而且更符合保持物体边缘的要求. 他们提出来的模型是求下列泛函的极小点 u:

$$E(u) = \frac{1}{2} \int_\Omega |u_d - Ku|^2\, dx + \lambda \int_\Omega \phi(|\nabla u|)\, dx, \tag{20.3}$$

其中 $\phi : \mathbb{R}^+ \to \mathbb{R}^-$ 是一个严格凸的、单调不减函数, 满足

$$\phi(0) = 0 \tag{20.4}$$

和

$$\exists c > 0, \exists b \geqslant 0, \text{使得}\, cs - b \leqslant \phi(s) \leqslant cs + b, \forall s. \tag{20.5}$$

然而在BV 函数空间上, 应当怎样理解 (20.3) 式中的第二个积分?

定理 20.3 (E. De Giorgi, L. Ambrosio, G. Buttazzo) 在 $\mathrm{BV}(\Omega)$ 上泛函 E 的松弛泛函是

$$\begin{aligned} RE(u) = &\frac{1}{2}\int_\Omega |u_d - Ku|^2 dx + \lambda \int_\Omega \phi(|\nabla u|)\, dx \\ &+\lambda c \int_{S_u} (u^+ - u^-)\, d\mathcal{H}^1 + \lambda c \int_{\Omega \setminus S_u} |C_u| dx, \end{aligned} \tag{20.6}$$

其中

$$Du = \nabla u \mathcal{L}^2 + (u^+ - u^-) n_u \mathcal{H}^1|_{S_u} + C_u$$

是 Lebesgue 分解. 参看 [Bu].

定理 20.4 在前面的假设下, 又设 $K \in \mathcal{L}(L^2, L^2)$, 且 $\|K \cdot 1\| \neq 0$, 则泛函 $RE(u)$ 有极小点 $u_0 \in \mathrm{BV}(\Omega)$.

证明 只要证明 RE 在 $\mathrm{BV}(\Omega)$ 上是强制的就够了. 设 $\{u_j\} \subset \mathrm{BV}(\Omega)$ 是 RE 的极小化序列. 由 (20.5) 与 (20.6), 可见 $\|Du_j\|(\Omega)$ 是有界的, 并且 $\|u_d - Ku_j\|_2$ 是有界的. 剩下来证明 $\|u_j\|_1$ 是有界的. 作分解

$$u_j = v_j + w_j,$$

其中 $w_j = \frac{1}{|\Omega|} \int_\Omega u_j\, dx$.

按 Poincaré 不等式 (简单性质 (5)) $\|v_j\|_2 \leqslant C\|Du_j\|(\Omega)$, 只需再证 $|w_j|$ 有界就够了. 令 $\|K \cdot 1\| = r, \|K\| = k_0$, 则由

$$\|u_d - Ku_j\|_2^2 \leqslant M,$$

可见

$$r|w_j|\,(r|w_j| - 2(k_0\|v_j\|_2 + \|u_d\|_2)) \leqslant (\|Kv_j - u_d\|_2 - |w_j| r)^2 \leqslant \|u_d - Ku_j\|_2^2 \leqslant M.$$

注意到已知 $\|v_j\|_2$ 是有界的, 所以 $|w_j|$ 是有界的.

利用松弛泛函 RE 的 *弱序列下半连续性, 所以存在极小点 $u_0 \in \mathrm{BV}(\Omega)$. □

参考文献

[CL] 张恭庆, 林源渠. 泛函分析讲义: 上册. 北京: 北京大学出版社, 1987.

[LL] M. A. 拉弗林契叶夫, L. A. 留斯切尔涅克著. 变分学教程. 曾鼎龢, 邓汉英, 王梓坤, 译. 北京: 高等教育出版社, 1955.

[Ka] 加藤敏夫. 变分学及其应用. 周怀生, 译. 上海: 上海科学技术出版社, 1961.

[Ad] R. A. Adams. Sobolev Spaces. Acad. Press, 1975 (中译本: 叶其孝、王耀东、应隆安、韩厚德、吴兰成, 译. 北京: 人民教育出版社, 1983).

[AFP] L. Ambrosio, N. Fusco, D. Pallara. Functions of Bounded Variation and Free Discontinuous Problems. Clarendon Press, Oxford, 2000.

[AE] J. P. Aubin, I. Ekeland. Applied Nonlinear Analysis. A John Wiley-Interscience Publication, 1984.

[AK] G. Aubert, P. Kornprobst. Mathematical Problems in Image Processing, Partial Differential Equations and the Calculus of Variations. Applied Math. Sciences 147, Springer-Verlag, 2002.

[BM] U. Brechtken, U. Manderschild. Introduction to the Calculus of Variations. Chapman Hall, 1991.

[Bu] G. Buttazzo. Semicontinuity, Relaxation and Integral Representation in the Calculus of Variations. Pitman Research Notes in Mathematics 207, Longman Scientific and Technical, 1989.

[BGH] G. Buttazzo, M. Giaquinta, S. Hildebrandt. One-dimensional Variational Problems, An Introduction. Clarendon Press, Oxford, 1998.

[Ch] K. C. Chang. Infinite Dimensional Morse Theory and Multiple Solution Prob-

lems, Birkhauser, 1993.

[Co] R. Courant. Dirichlet's Principle, Conformal Mapping, and Minimal Surfaces. New York, Interscience, 1950.

[Da] B. Dacorogna. Direct Methods in the Calculus of Variations. Applied Math. Sciences, 78, Springer-Verlag, 1989.

[Ek] I. Ekeland. Non-convex Minimization Problems. Bull. Amer. Math. Soc., 1979: 443-474.

[Fe] H. Federer. Geometric Measure Theory. Springer-Verlag, 1969.

[GF] I. M. Gelfand, S. V. Fomin. Calculus of Variations (English translated by R.A. Silverman). Prentice Hall, 1964.

[Gi] M. Giaquinta. The Regularity Problem of Extremals of Variational Integrals. Proc. NATO/LMS Advance Study Inst. on "Systems of Nonlinear Partial Differential Equations". Reidel Publ. Co., 1983.

[GH] M. Giaquinta, S. Hildebrandt. Calculus of Variations I The Lagrangian formalism. Grundlehren der mathematischen Wissenschaften, 310, Springer-Verlag, 1996.

[GT] D. Gilbarg, N. Trudinger. Elliptic Partial Differential Equations of Second Order, 2nd ed. Grundlehren der mathematischen Wissenschaften, 224, Springer-Verlag, 1983.

[Ha] R. Hardt (Ed.). Six Theorems in Variations. Student Math. Lib., 26, AMS, 2004.

[Hi1] D. Hilbert. Uber das Dirichletsche Prinzip. Jber. Deutsch Math. Vere., 8, 1900: 184-188.

[Hi] S. Hildebrandt. Boundary Behavior of Minimal Surfaces. Arch. Rat. Mech. Anal., 35, 1969: 47-82.

[HT] H. Hofer, J. F. Toland. Homoclinic, Heteroclinic and Periodic Solutions for Indefinite Hamiltonian Systems. Math. Ann., 268, 1984: 387-403.

[Hu] T. Hughes. Finite Element Method, Linear Static and Dynamic Finite Element Analysis. Dover Publications, 2000.

[Jo] C. Johnson. Numerical Solution of Partial Differential Equations by the Finite Element Method. Studentlitteratur, Lund, 1987.

[JJ] J. Jost, X Li-Jost. Calculus of Variations. Cambridge University Press, 1998.

[Jo] J. Jost. Two Dimensional Geometric Variational Problems. John Wiley and Sons, 1990.

[Le] G. Leitmann. The Calculus of Variations and Optimal Control, An Introduction. Plenum Press, 1981.

[Mac] C. R. MacCluer. Calculus of Variations, Mechanics, Control, and Other Appli-

cations. Pearson Education Ltd., 2005.

[MT] P. Mallianvin, A. Thalmaier. Stohastic Calculus of Variations in Mathematical Finance. Springer Finance. Springer-Verlag, 2006.

[Maw] J. Mawhin. Global Results for the Forced Pendulum Equation. Handbook of Differential Equations, Ordinary Differential Equations, (Ed. by A. Canada, P. Drabek, F. Fonda), Vol.1: 533-589.

[MW] J. Mawhin, M. Willem. Critical Point Theory and Hamiltonian Systems. Applied Math. Sciences, 74, Springer-Verlag, 1989.

[MS] J. Macki, A. Strauss. Introduction to Optimal Control Theory. Springer-Verlag, 1982.

[Mor] C. B. Morrey. Jr. Multiple Integrals in the Calculus of Variations. Springer-Verlag, 1966.

[Mos] J. Moser. Selected Chapters in the Calculus of Variations. Birkauser, 2003.

[Ne] Z. Nehari. Characteristic Values Associated with a Class of Nonlinear Second Order Equations. Acta Math., 105, 1961: 141-175.

[Os] R. Osserman. A survey of Minimal Surfaces. Dover Publications, 1986.

[PBGM] L. Pontryagin, V. Boltyanskii, R. Gamkrelidze, E. Mishchenko. The Mathematical Theory of Optimal Processes. Interscience Publishers, John Wiley and Sons, 1962.

[Po] S. Pohozaev. Nonlinear Variational Problems via the Fibering Method. Handbook of Differential Equations, Stationary Partial Differential Equations (Ed. by M. Chipot), Vol. 5, 2008: 49-209.

[Ra1] P. Rabinowitz. Minimax Methods in Critical Point Theory with Applications to Differential Equations. CBMS Regional Conference Series Math., AMS, 65, 1986.

[Ra2] P. Rabinowitz. Homoclinic Orbits for a Class of Hamiltonian Systems. Proc. Royal Soc. Edinburgh, 114 A, 1990: 33-38.

[Ra3] P. Rabinowitz. Periodic and Heteroclinic Solutions for a Periodic Hamiltonian System. Ann. Inst. Henri Poincare, Anal. Nonlineaire 6, 1989: 331-346.

[St1] M. Struwe. Variational Methods, Applications to Nonlinear Partial Differential Equations and Hamiltonian Systems. Springer-Verlag, 1990.

[St2] M. Struwe. Plateau's Problem and the Calculus of Variations, Mathematical Notes. Princeton University Press, 1988.

参考文献说明.

第一部分 (第一讲—第八讲) 主要参考:

[BGH], [LL], [GF], [GH], [JJ], [Ka], [Ma], [Mos].

第二部分 (第九讲—第十四讲) 主要参考:

[Ad], [AE], [CL], [Da], [Ek], [Gi], [JJ], [Mos], [Mor], [Ne], [Po].

第三部分 (第十五讲—第二十讲),

第十五讲参考: [Ch], [JJ], [Ra1], [St1], [MW].

第十六讲参考: [BGH], [HT], [Maw], [Mos], [Ra2], [Ra3].

第十七讲参考: [Jo], [Ha], [Hi], [Mor], [Os], [St2].

第十八讲参考: [Hu], [Jo].

第十九讲参考: [BM], [Ma], [Le], [PBGM].

第二十讲参考: [AFP], [AK], [BGH], [Bu].

索引

B

C

D